

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 14:00:34 ; Search time 194 Seconds
(without alignments)
513.092 Million cell updates/sec

Title: US-09-820-003C-2
Perfect score: 886
Sequence: 1 MSNMNPEYDYLFLKLLIGDS.....EKSVMKIQSTPVKSGSGGCC 173

Scoring table: BIOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	860	97.1	205	1	RBIA CANFA
2	860	97.1	205	1	RBIA_HUMAN
3	860	97.1	205	1	RBIA_MOUSE
4	860	97.1	205	1	RBIA_MOUSE
5	860	97.1	205	2	AAH66662
6	860	97.1	205	2	CAE11872
7	860	97.1	205	2	AAH62594
8	860	97.1	205	2	BAC28697
9	853	96.3	253	2	Q6ZPF0
10	853	96.3	253	2	BAC98287
11	831.5	93.8	204	2	Q6GLH0
12	831.5	93.8	204	2	Q7ZXF7
13	828	93.5	202	1	RBIA DISOM
14	800.5	90.3	201	2	Q7ZS50
15	800.5	90.3	201	2	AAH62857
16	799.5	90.2	201	2	Q7ZUV5
17	766.5	86.5	201	1	RBIB_HUMAN
18	763.5	86.2	201	1	Q6F1G4
19	760.5	85.8	201	1	RBIB_RAT
20	759.5	85.7	201	1	RBIB_MOUSE
21	752	84.9	301	2	Q7TPK7
22	733	82.7	205	1	RABI_LYMSM
23	683	77.1	205	2	O18133
24	671	75.7	202	2	O16143
25	670	75.6	164	2	Q610G4
26	670	75.6	164	2	AAH71442
27	660	74.5	205	2	Q9UAG6
28	658.5	74.3	203	1	YPT1_NEUCR
29	655.5	74.0	201	2	Q9HET3
30	653.5	73.8	201	2	Q9P8J7
31	646.5	73.0	203	1	YPT1_CHLRE

32	646.5	73.0	203	1	YPT1_SCHPO	P1620 schizosacch
33	644	72.7	202	2	Q40203	Q40203 lotus japon
34	642.5	72.5	203	1	YPT1_VOLCA	P1584 volvox cart
35	642	72.5	202	2	Q9SEH3	Q9SEH3 arabidopsis
36	640.5	72.3	212	2	Q7Q874	Q7Q874 anopheles g
37	640	72.2	202	2	Q9HDT5	Q9HDT5 trichoderma
38	640	72.2	203	2	Q6C9V1	Q6C9V1 yarrowia li
39	636.5	71.8	203	2	Q08154	Q08154 plasm sativ
40	635	71.8	243	2	Q7RVH8	Q7RVH8 neurospora
41	634.5	71.6	202	2	Q8K4S8	Q8K4S8 arabidopsis
42	633	71.4	202	2	Q9FPV4	Q9FPV4 arabidopsis
43	632.5	71.4	203	2	Q24112	Q24112 nicotiana p
44	632	71.3	202	2	Q9SXT5	Q9SXT5 cicier arlet
45	631.5	71.3	201	1	YPT1_PHYIN	Q01890 phycophthor

ALIGNMENTS

RESULT 1
RBIA CANFA STANDARD; PRT; 205 AA.
ID RBIA CANFA
AC P62822; P05711;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Ras-related protein Rab-1A.
GN Name=RBIA; Synonyms=RAB1;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cocker spaniel;
RX MEDLINE=91061765; PubMed=2121294;
RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;
RT "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell
line.";
RL Mol. Cell. Biol. 10:6578-6585 (1990).
RN [2]
RP ISOPRENOLID.
RX MEDLINE=91296801; PubMed=1648736;
RA Khosravi-Far R., Lutz R.J., Cox A.D., Conroy L., Bourne J.R.,
RA Sinensky M., Balch W.E., Buss J.B., Der C.J.;
RT "Isoprenoid modification of Rab proteins terminating in CC or CXC
motifs.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:6264-6268 (1991).
CC -1- FUNCTION: Probably required for transit of protein from the ER
through Golgi compartment. Binds GTP and GDP and possesses
intrinsic GTPase activity.
CC -1- SUBCELLULAR LOCATION: Golgi.
CC -1- PTM: Phosphorylated by CDC2 kinase during mitosis (By similarity).
CC -1- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; X56384; CAB56775.1; -.
DR HSSP; P07560; 1G16.
DR InterPro; IPR001806; Ras trnsfmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; Ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR TIGRFAMs; TIGR00231; small_gtp; 1.
DR Endoplasmic reticulum; Golgi stack; GTP-binding; Lipoprotein;
KW Prenylation; Protein transport.
FT NP_BIND 18 25 GTP (By similarity).

FT NP BIND 66 70 GTP (By similarity).
 FT NP BIND 124 127 Effector region (By similarity).
 FT DOMAIN 40 48 S-geranylgeranyl cysteine.
 FT LIPID 204 204 S-geranylgeranyl cysteine.
 FT LIPID 205 205 S-geranylgeranyl cysteine.
 SQ SEQUENCE 205 AA; 22678 MW; B2A8FA83B0FB17D6 CRC64;

Query Match 97.1%; Score 860; DB 1; Length 205;
 Best Local Similarity 84.4%; Pred. No. 2e-62;
 Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSNNPEVDYLFKLLIGDSGVGKSCLLRFADPTTSTYSTISTIGVDPKIRITIELDGKTI 60
 DB 1 MSSNNPEVDYLFKLLIGDSGVGKSCLLRFADPTTSTYSTISTIGVDPKIRITIELDGKTI 60
 QY 61 KQI-----ESPNNVKQWLQSIDRYASENVKL 88
 DB 61 KQIWDPAQGERFRTTSSYRGAGHIIWVDVTDQESFNNVKQWLQSIDRYASENVKL 120
 QY 89 LVGNKCDLITTKKVVDYTTAKEPADSLGIPLETSAKNTNVEQSPMTAAEIKKRMGPGA 148
 DB 121 LVGNKCDLITTKKVVDYTTAKEPADSLGIPLETSAKNTNVEQSPMTAAEIKKRMGPGA 180
 QY 149 TAGGAEKSNVKIQSTPPVQSGGGCC 173
 DB 181 TAGGAEKSNVKIQSTPPVQSGGGCC 205

RESULT 2
 ID RBIA HUMAN STANDARD; PRT; 205 AA.
 AC P62820; P11476; Q96N61; Q9Y3T2;
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Ras-related protein Rab-1a (YPT1-related protein).
 GN Name=RBIA; Synonyms=RB1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=89308668; PubMed=2501306;
 RA Zairacou A., Touchot N., Chardin P., Tavilian A.;
 RT "The human Rab genes encode a family of GTP-binding proteins related
 to yeast YPT1 and SEC4 products involved in secretion.";
 RL J. Biol. Chem. 264:12394-12401 (1989).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC TISSUE=Brain;
 RX MEDLINE=21154917; PubMed=11230166; DOI=10.1101/gr.154701;
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glaesl S.,
 RA Ansorge W., Boecker M., Bioecker H., Bauersachs S., Blum H.,
 RA Laber J., Duesterhoeft A., Beyer A., Koerner K., Strack N.,
 RA Mewes H.-W., Oettermann B., Obermaier B., Tampe D., Heudner D.,
 RA Wandt R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 analysis of 500 novel complete protein coding human cDNAs.";
 RL Genome Res. 11:422-435 (2001).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Colon;
 RA Koehler K., Beyer A., Mewes H.-W., Weil B., Amid C., Oeanger A.,
 RA Fodo G., Han M., Wiemann S.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu N., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,

RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shitatori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Nimomiyama K., Iehashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotta T.,
 RA Kusano J., Kanehori K., Takahashi-Fuji A., Hara H., Tanase T.-O.,
 RA Kusanu Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Imose N., Miasashiro K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shibata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Waki H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Omori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 cDNAs.";
 RL Nat. Genet. 36:40-45 (2004).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;
 RA Publ H.L. III, Ikeda S.R., Aronstein R.S.;
 RT "cDNA clones of human proteins involved in signal transduction
 sequenced by the Guthrie cDNA resource center (www.cdna.org).";
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain, and Placenta;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Ushed T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [7]
 RP PHOSPHORYLATION BY CDC2.
 RX MEDLINE=91218852; PubMed=1902553;
 RA Bailey E., Mc Caffrey M., Touchot N., Zairacou A., Goud B., Bornens M.;
 RT "Phosphorylation of two small GTP-binding proteins of the Rab family
 by p34cdc2.";
 RL Nature 350:715-718 (1991).
 RN [8]
 RP FUNCTION: Probably required for transit of protein from the ER
 through Golgi compartment. Binds GTP and GDP and possesses
 intrinsic GTPase activity.
 CC -!- SUBCELLULAR LOCATION: Golgi.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=P62820-1, P11476-1; Sequence=Displayed;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 14:08:40 ; Search time 38 seconds

(without alignments)
438.039 Million cell updates/sec

Title: US-09-820-003C-2

Perfect score: 886

Sequence: 1 MSSMNPEDYLFKLLIGDS.....EKNVKTQSTPVKSGGGCC 173

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	860	97.1	205	1	TVDGRP
2	860	97.1	205	1	TVHUP
3	860	97.1	205	1	TVMSYP
4	849	95.8	205	1	TVRTYP
5	812.5	91.7	201	2	D38625
6	771.5	87.1	201	2	S06147
7	733	82.7	205	2	S38339
8	671	75.7	202	2	JB0318
9	660	74.5	203	2	S30096
10	658.5	74.3	203	2	S30096
11	646.5	73.0	203	2	YC105
12	646.5	73.0	203	2	YC105
13	642.5	72.5	203	2	YC1247
14	631.5	71.3	201	2	UC5337
15	627.5	70.8	203	2	S34253
16	623.5	70.4	206	2	S04590
17	619.5	69.9	258	2	B86153
18	619	69.9	202	2	S41430
19	618.5	69.8	203	2	B38202
20	613.5	69.2	202	2	S38740
21	612	69.1	221	2	H71444
22	598	67.5	218	2	T07609
23	590.5	66.6	202	2	S72515
24	587	66.3	201	2	S39565
25	574	64.8	206	2	T14391
26	572.5	64.6	196	2	S0279
27	571	64.4	196	1	TVBYQ2
28	523	59.0	208	2	A38202
29	468.5	52.9	208	2	A34716

30	461	52.0	203	2	B34716	GTP-binding protei
31	447	50.5	216	2	T48378	GTP-binding protei
32	442	49.9	215	2	T14565	GTP-binding protei
33	439.5	49.6	215	2	S57478	GTP-binding protei
34	438.5	49.5	215	2	US0640	GTP-binding protei
35	435.5	49.2	200	2	S12790	GTP-binding protei
36	430.5	48.6	215	2	S57462	GTP-binding protei
37	429.5	48.5	216	2	S57471	GTP-binding protei
38	429.5	48.5	222	2	T14405	small GTP-binding
39	428.5	48.4	215	2	T45901	GTPase ARAB8 - Ar
40	426.5	48.1	215	2	S57474	GTP-binding protei
41	425.5	48.0	216	2	S33900	GTP-binding protei
42	418	47.2	207	2	B49647	GTP-binding protei
43	418	47.2	207	2	B36364	GTP-binding protei
44	410.5	46.3	206	2	T78851	GTP-binding protei
45	409.5	46.2	209	2	B38625	GTP-binding protei

ALIGNMENTS

RESULT 1

TVDGRP

GTP-binding protein Rabi - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 19-Feb-1994 #sequence revision 06-Dec-1996 #text_change 19-Jan-2001

C:Accession: S19104; A36364; S15600

R:Zerial, M.

submitted to the EMBL Data Library, August 1990

A:Reference number: S19104

A:Accession: S19104

A:Molecule type: mRNA

A:Residues: 1-205 <ZER>

A:Cross-references: EMBL:X56384

R:Chavrier, P.; Vingron, M.; Sander, C.; Simons, K.; Zerial, M.

Mol. Cell. Biol. 10, 6578-6585, 1990

A:Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.

A:Reference number: A36364; MUID:91061765; PMID:2123294

A:Accession: A36364

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-157, 'EK', 160-205 <CHA>

A:Cross-references: GB:X56384; NID:9913

C:Function:

A:Description: probably involved in protein transport from the endoplasmic reticulum thr

C:Superfamily: ras transforming protein; translation elongation factor Tu homolog

C:Keywords: endoplasmic reticulum; Golgi apparatus; GTP binding; lipidprotein; membrane p

tein

F:12-127/Domains: translation elongation factor Tu homology <RTU>

F:18-25/Region: nucleotide-binding motif A (P-loop)

F:62-67/Region: nucleotide-binding motif B

F:124-127/Region: GTP-binding NKXD motif

F:154-156/Region: GTP-binding SAK/L motif

F:24-25,43,124,125,127,154/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #stat

F:194/Binding site: phosphate (Ser) (covalent) (by cdc2 kinase) #statue predicted

F:204,205/Binding site: geranyl-geranyl (Cys) (covalent) #statue predicted

Query Match 97.1%; Score 860; DB 1; Length 205;

Best local similarity 84.4%; Pred. No. 1.7e-61;

Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSMNPEDYLFKLLIGDSGVKSCLLAFADPTYESYITGVDVFKRTIELDQKI 60

DB 1 MSSMNPEDYLFKLLIGDSGVKSCLLAFADPTYESYITGVDVFKRTIELDQKI 60

QY 61 KLOI-----ESFNNVQWLOEIDRYASENNKL 88

DB 61 KLOIWDTAGGERPRTTSSYRGAGHGIIVYVDYDQESFNNVQWLOEIDRYASENNKL 120

QY 89 LVGNKCDLTKKVVDDYTTAKEPADSLGIPLETSKAKNAIVNVEGSEFMTAAEIKRNGPGA 148

DB 121 LVGNKCDLTKKVVDDYTTAKEPADSLGIPLETSKAKNAIVNVEGSEFMTAAEIKRNGPGA 180

QY 149 TAGGAEKSNVVKIQSTPVKQSGGGCC 173
 Db 181 TAGGAEKSNVVKIQSTPVKQSGGGCC 205

RESULT 2

TVHUYF
 GTP-binding protein Rab1 - human
 N/Alternate names: protein DKFZ564B163.1
 C/Species: Homo sapiens (man)
 C/Date: 29-Jun-1990 #sequence_revision 06-Dec-1996 #text_change 09-Jul-2004
 C/Accession: A34323; T08698
 R/Zahraoui, A.; Touchot, N.; Chardin, P.; Tavittian, A.
 J. Biol. Chem. 264, 12394-12401, 1989
 A/Title: The human Rab gene encode a family of GTP-binding proteins related to yeast YF
 A/Reference number: A34323; MUID:89308668; PMID:2501306
 A/Accession: A34323
 A/Molecule type: mRNA
 A/Residues: 1-205 <ZAH>
 A/Cross-references: UNIPROT:P11476; GB:J04941; GB:M28209; NID:g55059; PIDN:AAA60240.1;
 A/Experimental source: pHeochromocytoma
 R/Habdruck, R.; Heubner, D.; Wewes, H.W.; Gassenhuber, J.; Wiemann, S.
 Submitted to the Protein Sequence Database, March 1999
 A/Reference number: 216471
 A/Accession: T08698
 A/Molecule type: mRNA
 A/Residues: 1-64,141-205 <WAW>
 A/Cross-references: EMBL:AL050268
 A/Experimental source: fetal brain; clone DKFZ564B163
 C/Genetics:
 A/Gene: GDB:RAB1
 A/Cross-references: GDB:118857; OMIM:179508
 A/Map position: 4p15.31-4p15.31
 A/Note: DKFZ564B163.1
 C/Function:
 A/Description: probably involved in protein transport from the endoplasmic reticulum th
 C/Superfamily: ras transforming protein; translation elongation factor Tu homology
 C/Keywords: endoplasmic reticulum; Golgi apparatus; GTP binding; lipoprotein; membrane p
 tein
 F:1-205/Product: GTP-binding protein Rab1. #status predicted <MAT1>
 F:1-64,141-205/Product: GTP-binding protein Rab1, splice variant #status predicted <MAT2
 F:12-127/Domains: translation elongation factor Tu homology <ETU>
 F:18-25/Region: nucleotide-binding motif A (P-loop)
 F:62-67/Region: nucleotide-binding motif B
 F:124-127/Region: GTP-binding NAKD motif
 F:154-156/Region: GTP-binding SAK/L motif
 F:24,25,43,124,125,127,154/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta
 F:194/Binding site: phosphate (Ser) (covalent) (by cdcd kinase) #status predicted
 F:204,205/Binding site: geranyl-geranyl (Cys) (covalent) #status experimental

Query Match 97.1%; Score 860; DB 1; Length 205;
 Best Local Similarity 84.4%; Pred. No. 1.7e-61;
 Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSMNEEYVLFKLLIGDSGVKSCLLRFADDTYESYISTIGVDFKIRITELDGKTI 60
 Db 1 MSSMNEEYVLFKLLIGDSGVKSCLLRFADDTYESYISTIGVDFKIRITELDGKTI 60
 QY 61 KLOI-----ESNNYKQWLOEIDRYASENVNKL 88
 Db 61 KLOIWDTAGOERFRITTSYRGAGIIIVYDVTDQESFNNVQWLOEIDRYASENVNKL 120
 QY 89 LVGNKCDLTKKVVDDTTAKEPADSLGIPLETSAKATNVQSFMTMAAEIKKMGPGA 148
 Db 121 LVGNKCDLTKKVVDDTTAKEPADSLGIPLETSAKATNVQSFMTMAAEIKKMGPGA 180
 QY 149 TAGGAEKSNVVKIQSTPVKQSGGGCC 173
 Db 181 TAGGAEKSNVVKIQSTPVKQSGGGCC 205

RESULT 3
 TVMSYP

GTP-binding protein ypt1 - mouse
 N/Alternate names: GTP-binding protein Rab1; ras-related protein ypt1; transforming prote
 C/Species: Mus musculus (house mouse)
 C/Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
 C/Accession: S05551; S06285
 R/Wichmann, H.; Disela, C.; Haubruck, H.; Gallwitz, D.
 Nucleic Acids Res. 17, 6737-6738, 1989
 A/Title: Nucleotide sequence of the mouse ypt1 gene encoding a ras-related GTP-binding p
 A/Reference number: S05551; MUID:89386011; PMID:2506528
 A/Accession: S05551
 A/Molecule type: DNA
 A/Residues: 1-205 <WIC>
 A/Cross-references: UNIPROT:P11476; EMBL:X15744; NID:g55458; PIDN:CAA33760.1; PID:g76315f
 R/Habdruck, H.; Disela, C.; Wagner, P.; Gallwitz, D.
 EMBO J. 6, 4049-4053, 1987
 A/Title: The ras-related ypt protein is an ubiquitous eukaryotic protein: isolation and
 A/Reference number: S06285; MUID:8816649; PMID:3127202
 A/Accession: S06285
 A/Molecule type: mRNA
 A/Residues: 1-205 <HAD>
 C/Genetics:
 A/Gene: ypt1
 A/Intons: 8/2; 32/3; 64/3; 96/3; 140/3
 C/Function:
 A/Description: probably involved in protein transport from the endoplasmic reticulum thr
 C/Superfamily: ras transforming protein; translation elongation factor Tu homology
 C/Keywords: endoplasmic reticulum; Golgi apparatus; GTP binding; lipoprotein; membrane p
 tein

F:12-127/Domains: translation elongation factor Tu homology <ETU>
 F:18-25/Region: nucleotide-binding motif A (P-loop)
 F:62-67/Region: nucleotide-binding motif B
 F:124-127/Region: GTP-binding NAKD motif
 F:154-156/Region: GTP-binding SAK/L motif
 F:24,25,43,124,125,127,154/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta
 F:194/Binding site: phosphate (Ser) (covalent) (by cdcd kinase) #status predicted
 F:204,205/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 97.1%; Score 860; DB 1; Length 205;
 Best Local Similarity 84.4%; Pred. No. 1.7e-61;
 Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSMNEEYVLFKLLIGDSGVKSCLLRFADDTYESYISTIGVDFKIRITELDGKTI 60
 Db 1 MSSMNEEYVLFKLLIGDSGVKSCLLRFADDTYESYISTIGVDFKIRITELDGKTI 60
 QY 61 KLOI-----ESNNYKQWLOEIDRYASENVNKL 88
 Db 61 KLOIWDTAGOERFRITTSYRGAGIIIVYDVTDQESFNNVQWLOEIDRYASENVNKL 120
 QY 89 LVGNKCDLTKKVVDDTTAKEPADSLGIPLETSAKATNVQSFMTMAAEIKKMGPGA 148
 Db 121 LVGNKCDLTKKVVDDTTAKEPADSLGIPLETSAKATNVQSFMTMAAEIKKMGPGA 180
 QY 149 TAGGAEKSNVVKIQSTPVKQSGGGCC 173
 Db 181 TAGGAEKSNVVKIQSTPVKQSGGGCC 205

RESULT 4

TVRYP
 GTP-binding protein Rab1 - rat
 N/Alternate names: transforming protein ypt1 homolog
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 03-Aug-1992 #sequence_revision 06-Dec-1996 #text_change 09-Jul-2004
 C/Accession: A39963
 R/Zahraoui, A.; Touchot, N.; Chardin, P.; Tavittian, A.
 Proc. Natl. Acad. Sci. U.S.A. 84, 8210-8214, 1987
 A/Title: Four additional members of the ras gene superfamily isolated by an oligonucleot
 A/Reference number: A39963; MUID:88068563; PMID:3317403
 A/Accession: A39963
 A/Molecule type: mRNA
 A/Residues: 1-205 <TOU>
 A/Cross-references: UNIPROT:P05711; GB:J02998; NID:g206552; PIDN:AAA42006.1; PID:g206553

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 14:08:55 ; Search time 40 Seconds
(without alignments)
286.825 Million cell updates/sec

Title: US-09-820-003C-2

Sequence: 1 MSSMNEPYDYLFTLLIGDS.....EKSNVKIQSTPVKQSGGCC 173

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCUS COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	857	96.7	205	4 US-09-709-103-49	Sequence 49, App1
2	857	96.7	205	4 US-09-439-410A-49	Sequence 49, App1
3	813.5	91.8	201	2 US-08-531-525-13	Sequence 13, App1
4	813.5	91.8	201	2 US-08-718-270A-13	Sequence 13, App1
5	766.5	86.5	201	2 US-08-916-901-3	Sequence 3, App1
6	766.5	86.5	201	2 US-09-154-602-3	Sequence 3, App1
7	760.5	85.8	201	2 US-08-916-901-8	Sequence 8, App1
8	760.5	85.8	201	2 US-09-154-602-8	Sequence 8, App1
9	702.5	79.3	202	2 US-08-531-525-14	Sequence 14, App1
10	702.5	79.3	202	2 US-08-718-270A-14	Sequence 14, App1
11	596.5	67.3	227	4 US-09-248-796A-20291	Sequence 20291, A
12	439	49.5	215	2 US-08-531-525-10	Sequence 10, App1
13	439	49.5	215	2 US-08-718-270A-10	Sequence 10, App1
14	424	47.9	207	2 US-08-531-525-35	Sequence 35, App1
15	424	47.9	207	2 US-08-718-270A-35	Sequence 35, App1
16	418	47.2	207	2 US-08-824-873-4	Sequence 4, App1
17	418	47.2	207	2 US-09-198-184-4	Sequence 4, App1
18	417.5	47.1	205	2 US-08-531-525-25	Sequence 25, App1
19	417.5	47.1	205	2 US-08-718-270A-25	Sequence 25, App1
20	403	45.5	203	4 US-09-255-920A-12	Sequence 12, App1
21	390	44.0	198	2 US-08-531-525-51	Sequence 51, App1
22	390	44.0	198	2 US-08-718-270A-51	Sequence 51, App1
23	366	41.3	194	2 US-08-531-525-34	Sequence 34, App1
24	366	41.3	194	2 US-08-718-270A-34	Sequence 34, App1
25	360	40.6	213	4 US-09-270-767-46812	Sequence 46812, A
26	360	40.6	213	4 US-09-248-796A-20293	Sequence 20293, A
27	334.5	37.8	214	4 US-09-270-767-33012	Sequence 33012, A

28	334.5	37.8	214	4 US-09-270-767-48229	Sequence 48229, A
29	334	37.7	208	2 US-08-531-525-17	Sequence 17, App1
30	334	37.7	208	2 US-08-718-270A-17	Sequence 17, App1
31	333	37.6	212	3 US-09-399-913-67	Sequence 67, App1
32	333	37.6	212	3 US-09-350-614-67	Sequence 67, App1
33	330.5	37.3	218	4 US-08-531-525-15	Sequence 19, App1
34	330.5	37.3	218	4 US-08-718-270A-19	Sequence 19, App1
35	330	37.2	212	2 US-08-531-525-18	Sequence 18, App1
36	330	37.2	212	2 US-08-718-270A-18	Sequence 18, App1
37	325.5	36.7	217	4 US-09-255-920A-15	Sequence 15, App1
38	322	36.3	210	2 US-08-531-525-16	Sequence 16, App1
39	322	36.3	210	2 US-08-718-270A-16	Sequence 16, App1
40	319	36.0	214	2 US-08-531-525-52	Sequence 52, App1
41	319	36.0	214	2 US-08-718-270A-52	Sequence 52, App1
42	317	35.8	128	4 US-09-513-999C-5775	Sequence 5775, App1
43	316.5	35.7	190	2 US-08-824-873-3	Sequence 3, App1
44	316.5	35.7	190	3 US-09-198-184-3	Sequence 3, App1
45	316.5	35.7	190	4 US-09-255-920A-7	Sequence 7, App1

ALIGNMENTS

```

RESULT 1
US-09-709-103-49
; Sequence 49, Application US/09709103
; Patent No. 6733991
; GENERAL INFORMATION:
; APPLICANT: Cismowski, Mary
; TITLE OF INVENTION: AGS Proteins and Nucleic Acid Molecules and Uses Therefor
; FILE REFERENCE: 60388-A-PC-US
; CURRENT APPLICATION NUMBER: US/09/709,103
; CURRENT FILING DATE: 2000-11-08
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-709-103-49

Query Match      96.7%; Score 857; DB 4; Length 205;
Best Local Similarity 83.9%; Pred. No. 2, 2e-92;
Matches 172; Conservative 1; Mismatches 0; Indels 32; Gaps 1;

QY      1 MSSMNEPYDYLFTLLIGDSGVKSCLLRFADDTYTESYISTIGVDFKIRTIELDGKTI 60
      |||
      1 MSSMNEPYDYLFTLLIGDSGVKSCLLRFADDTYTESYISTIGVDFKIRTIELDGKTI 60
Db

QY      61 KLOI-----ESPNNKQWIOEIDRYASEVNNKL 88
      |||
      61 KQIOWTAGERFRITTSYRGAGHIIYVYDTDESFPNNVQWIOEIDRYASEVNNKL 120
Db

QY      89 LVNKKCDLTKKVVDTTAKFADSLGIPFIETSAKNATVVEOSFPTMAEIKRKGPGA 148
      |||
      121 LVNKKCDLTKKVVDTTAKFADSLGIPFIETSAKNATVVEOSFPTMAEIKRKGPGA 180
Db

QY      149 TAGGAEKSNVKIOSTPVKQSGGCC 173
      |||
      181 TAGGAEKSNVKIOSTPVKQSGGCC 205
Db

RESULT 2
US-09-439-410A-49
; Sequence 49, Application US/09439410A
; Patent No. 6746852
; GENERAL INFORMATION:
; APPLICANT: Cismowski, Mary
; TITLE OF INVENTION: AGS PROTEIN AND NUCLEOTIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCE: 1919/60388-B
; CURRENT APPLICATION NUMBER: US/09/439,410A

```

CURRENT FILING DATE: 1999-11-11
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn version 3.1
SEQ ID NO 49
LENGTH: 205
TYPE: PRT
ORGANISM: Homo sapiens
US-09-439-410A-49

Query Match 96.7%; Score 857; DB 4; Length 205;
Best Local Similarity 83.9%; Pred. No. 2,2e-92;
Matches 172; Conservative 1; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSNNPEYDYLKLLIGDSGVGKSCLLRPADDTYTESYSTIGVDFKRTIELDGKTI 60
DB 1 MSSNNPEYDYLKLLIGDSGVGKSCLLRPADDTYTESYSTIGVDFKRTIELDGKTI 60
QY 61 KLOI-----ESFNNVKOMLOEIDRYASENVNKL 88
DB 61 KLOIWDPAQGERFRITSSYRGAGHIIIVYDVTDOESFNNVKOMLOEIDRYASENVNKL 120
QY 89 LVNKKCDLTTKKVVDYTTAKEFADSLGIPLETSAKNAATNVEQSFMTMAEIKRMGPGA 148
DB 121 LVNKKCDLTTKKVVDYTTAKEFADSLGIPLETSAKNAATNVEQSFMTMAEIKRMGPGA 180
QY 149 TAGGAESKNVKIQSTPVKOSGGGCC 173
DB 181 TAGGAESKNVKIQSTPVKOSGGGCC 205

RESULT 3

US-08-531-525-13
Sequence 13, Application US/08531525
Patent No. 5840683

GENERAL INFORMATION:
APPLICANT: Hlawka, Joseph J.
APPLICANT: Pincus, Matthew R.
APPLICANT: No. 5840683le, John F.
APPLICANT: Abajian, Henry B.
APPLICANT: Kende, Andrew S.
TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
TITLE OF INVENTION: of p21 Ras
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Greenlee and Winner, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531,525
FILING DATE: 21-SEP-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Farber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 37-94
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Discopyge ommata
US-08-531-525-13

Query Match 91.8%; Score 813.5; DB 2; Length 201;
Best Local Similarity 81.2%; Pred. No. 2,7e-87;
Matches 164; Conservative 3; Mismatches 2; Indels 33; Gaps 2;

QY 4 MNEPYDYLKLLIGDSGVGKSCLLRPADDTYTESYSTIGVDFKRTIELDGKTIKQ 63
DB 1 MNEPYDYLKLLIGDSGVGKSCLLRPADDTYTESYSTIGVDFKRTIELDGKTIKQ 60
QY 64 I-----ESFNNVKOMLOEIDRYASENVNKL 91
DB 61 IWDPAQGERFRITSSYRGAGHIIIVYDVTDOESFNNVKOMLOEIDRYASENVNKL 120
QY 92 NKCDLTTKKVVDYTTAKEFADSLGIPLETSAKNAATNVEQSFMTMAEIKRMGPATAG 151
DB 121 NKCDLTTKKVVDYTTAKEFADSLGIPLETSAKNAATNVEQSFMTMAEIKRMGPATAG 179
QY 152 GAESKNVKIQSTPVKOSGGGCC 173
DB 180 GSEKSNVNIQSTPVKOSGGGCC 201

RESULT 4

US-08-718-270A-13
Sequence 13, Application US/08718270A
Patent No. 5910478

GENERAL INFORMATION:
APPLICANT: Hlawka, Joseph J.
APPLICANT: Pincus, Matthew R.
APPLICANT: No. 5910478le, John F.
APPLICANT: Abajian, Henry B.
APPLICANT: Kende, Andrew S.
TITLE OF INVENTION: Peptidomimetics Inhibiting
TITLE OF INVENTION: the Oncogenic Action of p21 Ras
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718,270A
FILING DATE: 20-SEP-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/531,525
FILING DATE: 21-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,091
FILING DATE: 21-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Farber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 78-95
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
STRANDEDNESS: single

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 13:59:49 ; Search time 156 Seconds
(without alignment)
397.822 Million cell updates/sec

Title: US-09-820-003C-2

Perfect score: 886
Sequence: 1 MSSNNPEYDYLFLKLLIGDS.....EKSNNVKIQSTPVKSGGGGCC 173

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729239 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 405 summaries

Database : A_Geneseq_23Sep04:*

1: _geneseqp1980s:*
2: _geneseqp1990s:*
3: _geneseqp2000s:*
4: _geneseqp2001s:*
5: _geneseqp2002s:*
6: _geneseqp2003as:*
7: _geneseqp2003bs:*
8: _geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	886	100.0	173	ABG72742	ABG72742 Human Ras
2	886	100.0	173	ADM87199	ADM87199 Human pro
3	860	97.1	205	AAB34843	AAB34843 Gene 44 h
4	860	97.1	205	AAB34844	AAB34844 Human sec
5	860	97.1	205	ADB61461	ADB61461 Amyloid b
6	860	97.1	205	ADB63416	ADB63416 Human pro
7	860	97.1	205	ADB63412	ADB63412 Human pro
8	859	97.0	205	ADB61441	ADB61441 205 amino
9	817	92.2	198	AAB34816	AAB34816 Human sec
10	792.5	89.4	169	AAE29159	AAE29159 Human Ras
11	768.5	86.7	221	AAO13525	AAO13525 Human pol
12	768.5	86.7	224	AAB58758	AAB58758 Breast an
13	766.5	86.5	201	AAV00919	AAV00919 Human Rab
14	766.5	86.5	201	AAU28024	AAU28024 Novel hum
15	766.5	86.5	201	ABU52701	ABU52701 Human int
16	766.5	86.5	201	ADF76657	ADF76657 Novel hum
17	766.5	86.5	201	ADJ69790	ADJ69790 Human hea
18	760.5	85.8	201	ABU52702	ABU52702 Intracell
19	760.5	85.8	201	ADB63414	ADB63414 Rat prote
20	760.5	85.8	201	ADB63410	ADB63410 Rat prote
21	760.5	85.8	201	ADB58450	ADB58450 Rat prote
22	760.5	85.8	201	ADB83481	ADB83481 Rat prote
23	760.5	85.8	201	ADD47400	ADD47400 Rat prote
24	700	79.0	141	ADA54719	ADA54719 Human pro
25	683	77.1	205	ABB59808	ABB59808 Drosophila

26	637	71.9	202	3	AAG07763	Aag07763 Arabidops
27	635	71.7	225	6	ABJ26183	Abj26183 Aspergill
28	633	71.4	202	3	AAG10858	Aag10858 Arabidops
29	631.5	71.3	203	8	ADG00360	Adg00360 Nicotiana
30	619.5	69.9	258	3	AAG30498	Aag30498 Arabidops
31	618.5	69.8	203	3	AAG30499	Aag30499 Arabidops
32	612	69.1	218	3	AAG30710	Aag30710 Arabidops
33	612	69.1	221	3	AAG30709	Aag30709 Arabidops
34	612	69.1	254	3	AAG45323	Aag45323 Arabidops
35	596.5	67.3	207	5	ABP73466	Abp73466 Candida a
36	580	65.5	197	3	AAG48820	Aag48820 Arabidops
37	571	64.4	206	6	ABR53078	AbR53078 Protein s
38	571	64.4	206	7	ADK62970	Adk62970 Disease t
39	513.5	58.0	153	8	ADP29898	Adp29898 Human sec
40	447	50.5	216	3	ADN72851	Adn72851 Thale cre
41	444	50.1	218	3	AAG19220	Aag19220 Arabidops
42	442	49.9	216	3	AAG08006	Aag08006 Arabidops
43	440	49.7	206	6	ABJ26542	Abj26542 Aspergill
44	440	49.7	206	6	ABJ25583	Abj25583 Aspergill
45	431	48.6	215	3	AAG35215	Aag35215 Zea mays

ALIGNMENTS

RESULT 1	
ABG72742	standard; protein; 173 AA.
ID	ABG72742
AC	ABG72742;
DT	19-FEB-2003 (first entry)
DE	Human Ras-like protein.
XX	Human, Ras-like protein; Ras; low molecular weight; LMW;
XX	small regulatory guanine nucleotide-binding protein; GTP-binding protein;
XX	small G protein; signal transduction; regulation; modulator; diagnosis;
XX	apoptosis; AIDS; neurodegenerative disease; Alzheimer's disease;
XX	Parkinson's disease; myelodysplastic syndrome; aplastic anemia;
XX	ischemic injury; myocardial infarction; stroke; toxin-induced disease;
XX	alcohol-induced liver damage; cirrhosis; wasting disease; cachexia;
XX	viral infection; hepatitis; osteoporosis; cell proliferation; cancer;
XX	leukemia; inflammation; allergy; asthma; atherosclerosis;
XX	diabetes mellitus; rheumatoid arthritis; therapeutic; transgenic;
XX	pharmacogenomic analysis; genotype; antisense; gene therapy.
XX	Homo sapiens.
XX	US2002142382-A1.
XX	03-OCT-2002.
XX	29-MAR-2001; 2001US-00820003.
XX	29-MAR-2001; 2001US-00820003.
XX	(MERK/) MERKULOV G. V.
XX	(DFRA/) DI FRANCESCO V.
XX	(BEAS/) BEASLEY E. M.
XX	Merkulov GV, Di Francesco V, Beasley EM;
XX	WPI; 2003-102518/09.
XX	N-PSDB; ABX13973.
XX	Novel isolated human Ras-like polypeptide useful for diagnosing,
XX	preventing and treating inflammation and disorders associated with cell
XX	proliferation and apoptosis.
XX	Claim 1; Fig 2; 70pp; English.
XX	The invention discloses an isolated human Ras-like polypeptide and the

CC polynucleotide that encodes it. Ras proteins are small (low molecular
 CC weight, LMW) regulatory guanine nucleotide-binding protein (GTP-binding
 CC proteins or small G proteins) and are key relays in the signal
 CC transduction cascade. The Ras-like proteins participate in a wide range
 CC of regulatory functions. The Ras-like protein is useful for identifying a
 CC modulator of function or activity and for identifying an agent that binds
 CC to it. The polypeptide and antibody are useful in the diagnosis,
 CC prevention and treatment of disorders associated with an increase in
 CC apoptosis, such as AIDS, neurodegenerative diseases, such as Alzheimer's
 CC disease, Parkinson's disease, myelodysplastic syndromes, such as aplastic
 CC anaemia, ischaemic injuries, such as myocardial infarction, stroke, toxin
 CC -induced diseases, such as alcohol-induced liver damage, cirrhosis,
 CC wasting diseases, such as cachexia, viral infections, such as hepatitis B
 CC and C and osteoporosis, cell proliferation disorders, such as cancer and
 CC leukemia and inflammation disorders, such as allergy, asthma,
 CC atherosclerosis, diabetes mellitus and rheumatoid arthritis. The
 CC polypeptide and polynucleotide are useful as models for the development
 CC of human therapeutics. The polypeptide is useful to raise antibodies or
 CC to elicit an immune response, to determine levels of the protein in
 CC biological fluids and tissues, in drug screening assays and in
 CC pharmacogenomic analysis. The polynucleotide is useful for constructing
 CC transgenic animals, for monitoring the effectiveness of modulating
 CC compounds, for testing an individual for a genotype, as antisense
 CC constructs, and for gene therapy. The sequence presented is the human Ras-
 CC like protein

XX Sequence 173 AA;

Query Match 100.0%; Score 886; DB 6; Length 173;
 Best Local Similarity 100.0%; Pred. No. 5,7e-81;

- Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSNNPEVDYLFKLLIGDSGVGKSCLLRFADPTYSYISTIGVDFKIRITLEDKRTI 60
 Db 1 MSSNNPEVDYLFKLLIGDSGVGKSCLLRFADPTYSYISTIGVDFKIRITLEDKRTI 60
 QY 61 KLOIESFNNVKKQWLOEIDRYASENVNKKLVGNKCDLTTKKVVDYTTAKFADSLGIPFLE 120
 Db 61 KLOIESFNNVKKQWLOEIDRYASENVNKKLVGNKCDLTTKKVVDYTTAKFADSLGIPFLE 120
 QY 121 TSAKNATNVEOSFMTMAAEIKKRMGPATAGAGKSNVKIOSTPVKSGGGCC 173
 Db 121 TSAKNATNVEOSFMTMAAEIKKRMGPATAGAGKSNVKIOSTPVKSGGGCC 173

RESULT 2

ADM87199 standard; protein; 173 AA.

ADM87199;

03-JUN-2004 (first entry)

Human protein SEQ ID NO:292.

respiratory; cytosolic; antiarthritic; antiinflammatory;
 gastrointestinal; antibacterial; immunosuppressive; antidiabetic;
 antirheumatic; gene therapy; molecular weight marker; chromosome marker;
 chromosome tag; genetic fingerprinting; nutritional supplement; cancer;
 inflammatory condition; arthritis; inflammatory bowel disease;
 Crohn's disease; sepsis; rheumatoid arthritis; diabetes mellitus type 1;
 graft versus host disease; human.

Homo sapiens.

WO2004009834-A2.

29-JAN-2004.

19-JUL-2002; 2002WO-US022858.

21-JUL-2001; 2001US-0306971P.
 28-MAR-2002; 2002US-00112944.

XX (NUVE-) NUVELO INC.
 PA Tang YI, Yang Y, Meng G, Zhang J, Ren F, Xue A, Wang J;
 XX Weinman T, Ghosh M, Wang D, Zhao Q, Wang Z;
 PI WPI; 2004-143291/14.
 DR N-PSDB; ADM66955.
 XX
 PS Claim 20; SEQ ID NO 292; 591dp; English.

CC The present invention describes an isolated polynucleotide (1): (a)
 CC comprising a nucleotide sequence selected from SEQ ID NO:1-244; or (b)
 CC which encodes a polypeptide with biological activity, where the
 CC polynucleotide hybridises to (i) under stringent hybridisation conditions
 CC or has greater than 99% sequence identity with (1). (1) has respiratory,
 CC cytosolic, antiarthritic, antiinflammatory, gastrointestinal,
 CC antibacterial, immunosuppressive, antidiabetic and antirheumatic
 CC activities, and can be used in gene therapy. (1) can be used for
 CC generating polynucleotides encoding chimeric or fusion proteins and
 CC heterologous protein sequences. The polynucleotides can be used to
 CC express recombinant protein for analysis, characterisation or therapeutic
 CC use, as markers for tissues in which the corresponding protein is
 CC preferentially expressed, as molecular weight markers on gels, as
 CC chromosome markers or tags to identify chromosomes or to map related gene
 CC positions; to compare with endogenous DNA sequences in patients to
 CC identify potential genetic disorders; as probes to hybridise and discover
 CC genes, related DNA sequences; as a source of information to derive PCR
 CC primers for genetic fingerprinting; as a probe to subtract-out known
 CC sequences in the process of discovering other novel polynucleotides; for
 CC selecting and making oligomers for attachment to a gene chip or other
 CC support, including for examination of expression patterns; to raise anti-
 CC protein antibodies using DNA immunisation techniques; and as an antigen
 CC to raise anti-DNA antibodies or elicit another immune response. The
 CC polynucleotides and polypeptides can also be used as nutritional sources
 CC or supplements, e.g. as a protein or amino acid supplement, as a carbon
 CC source, as a nitrogen source or as a source of carbohydrates. The
 CC polynucleotides and polypeptides can also be used treat cancer. The
 CC compositions are useful for promoting better or faster closure of non-
 CC healing wounds, for the generation and regeneration of tissues, for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, and conditions resulting from
 CC systemic cytokine damage. The compositions can also be used to treat
 CC inflammatory conditions (e.g. arthritis, inflammatory bowel disease or
 CC Crohn's disease), sepsis, rheumatoid arthritis, diabetes mellitus type 1
 CC or graft versus host disease. The present sequence represents a novel
 CC human polypeptide sequence from the present invention. N.B. The sequences
 CC for this patent were obtained from the USPTO web site from an equivalent
 CC US patent US20040048249A1.

XX Sequence 173 AA;

Query Match 100.0%; Score 886; DB 6; Length 173;
 Best Local Similarity 100.0%; Pred. No. 5,7e-81;

- Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSNNPEVDYLFKLLIGDSGVGKSCLLRFADPTYSYISTIGVDFKIRITLEDKRTI 60
 Db 1 MSSNNPEVDYLFKLLIGDSGVGKSCLLRFADPTYSYISTIGVDFKIRITLEDKRTI 60
 QY 61 KLOIESFNNVKKQWLOEIDRYASENVNKKLVGNKCDLTTKKVVDYTTAKFADSLGIPFLE 120
 Db 61 KLOIESFNNVKKQWLOEIDRYASENVNKKLVGNKCDLTTKKVVDYTTAKFADSLGIPFLE 120
 QY 121 TSAKNATNVEOSFMTMAAEIKKRMGPATAGAGKSNVKIOSTPVKSGGGCC 173
 Db 121 TSAKNATNVEOSFMTMAAEIKKRMGPATAGAGKSNVKIOSTPVKSGGGCC 173

Result No.	Score	Match	Query Length	DB	ID	Description
1	1201.2	85.5	1927	3	CR616184	CR616184 full-length
2	1188	84.6	1440	3	CR593530	CR593530 full-length
3	1041.4	74.1	2411	3	AK034408	AK034408 Mus musc
4	947.2	67.4	990	3	CR606852	CR606852 full-length
5	930.2	66.2	1307	6	CEP110882	CEP110882 Shulkromi
6	914	65.1	1091	3	CR603642	CR603642 full-length
7	905.2	64.4	992	5	BX424956	BX424956 BX424956
8	890	63.3	1004	5	BX332698	BX332698 BX332698
9	885.6	63.0	1013	5	BX345432	BX345432 BX345432
10	866.8	61.7	1005	5	BX422946	BX422946 BX422946
11	862.4	59.4	908	7	CK9898121	CK9898121 bx27506.h
12	834.2	59.4	1117	1	ALJ539022	ALJ539022 ALJ539022
13	825.4	58.7	900	5	BX345431	BX345431 BX345431
14	818.4	58.2	1120	5	BX440905	BX440905 BX440905
15	817.6	58.2	970	5	BU559999	BU559999 AGENCOUR
16	810.4	57.7	886	5	BX413207	BX413207 BX413207
17	777.6	55.3	811	7	CP455551	CP455551 AGENCOUR
18	761.2	54.2	788	5	BU927812	BU927812 AGENCOUR
19	754.8	53.7	864	7	CF619291	CF619291 AGENCOUR
20	754.4	53.7	1017	5	BX332699	BX332699 BX332699
21	752	53.5	784	5	BU626516	BU626516 UI-H-PTO
22	751.8	53.5	775	6	CA443454	CA443454 UI-H-DPO
23	735.6	52.1	921	6	CD514144	CD514144 AGENCOUR
24	732.6	52.1	1054	5	BX426888	BX426888 BX426888

C	25	733.4	52.1	777	5	BQ014597	UI-H-ED1-
C	26	729.2	51.9	1094	7	CK230528	ILLUMIGEN
C	27	727.6	51.8	811	1	AA746643	nm27c08.8
C	28	723.8	51.5	817	5	B1869965	603394116
C	29	723.2	51.5	1127	5	BA6363467	BA6363467
C	30	721.4	51.3	939	5	BX411609	BX411609
C	31	719	51.2	1047	7	CN642107	ILLUMIGEN
C	32	718.6	51.1	756	7	CN433400	BE030001B
C	33	717.8	51.1	730	5	BUE688425	UI-CF-EC1
C	34	714.6	50.9	994	7	CN647880	ILLUMIGEN
C	35	714.6	50.9	1101	5	BX448589	BX448589
C	36	713.4	50.8	737	5	BUE622372	UI-H-FH1-
C	37	712	50.7	764	5	BUE68906	UI-CF-PNO
C	38	708.8	50.4	910	5	BQ276678	AGENCOURT
C	39	708.4	50.4	761	5	BQ772048	UI-H-EZ1-
C	40	704.4	50.1	1105	5	BM452262	AGENCOURT
C	41	703.6	50.1	725	5	BQ448090	UI-H-EU1-
C	42	703.4	50.1	708	7	CN285912	ILL0005939
C	43	702	50.0	796	4	B1913092	UI-H-19695
C	44	696.6	49.6	752	6	CA446990	UI-H-ED0-
C	45	695.6	49.5	745	6	CB456571	CB456571 713844 MA

ALIGNMENTS

RESULT 1	
CR616184	
LOCUS	
DEFINITION	CR616184 1927 bp mRNA linear HTC 21-JUL-2004 (human).
ACCESSION	CR616184
VERSION	CR616184.1 GI:50496991
KEYWORDS	HTC; CNSLT CDNA.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1927) Li,W.B., Gruber,C., Jesse,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished Contact : Peng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/InvitroGen Corporation 1600 Faraday Avenue 2 (bases 1 to 1927) Genoscope. Direct Submission Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segret@genoscope.cns.fr Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES	
Source	Location/Qualifiers 1..1927 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSDDM010YN17" /tissue type="Fetal liver" /plasmid="pCMVSPORT_6"
ORIGIN	
Query Match	85.5%; Score 1201.2; DB 3; Length 1927;
Best Local Similarity	93.0%; Pred. No. 2,5e-198;
Matches 1320; Conservative	0; Mismatches 3; Indels 97; Gaps 2;

```

Oy      1 AAGGATAGCTGACT -CGGGCGGTGCTATTGTGTTCTTAGGGGACGGAATAAGGGAGA 59
Db      12 AAGGATAGCTGATGGCGGGCGGCTGCTGATTGTGTTCTTAGGGGACGGAATAAGGGAGA 71

```

QY	60	CGTTTGCTCCCGGGAACGCCATATCTCATTCCTTTCTTTCGATTAACCGTGCGCGAG	119
Db	72	CGTTTCTCTCCCGGAACGCCATATCTCATTCCTTTCTTTCGATTAACCGTGCGCGAG	121
QY	120	AGTCAGGCGCGCGGCTGCGGCAGCAAGGCGCGGTGCGCGCGCGGAGCTCAGTGAC	179
Db	132	AGTCAGGCGCGCGGCTGCGGCAGCAAGGCGCGGTGCGCGCGCGGAGCTCAGTGAC	191
QY	180	ATGTCAGAGATGAATCCCGAATATGATTAATTAATTCAGATTACTTCTGATTGCGACTCA	239
Db	192	ATGTCAGAGATGAATCCCGAATATGATTAATTAATTCAGATTACTTCTGATTGCGACTCA	251
QY	240	GGGGTTGAAAGCTTGCCTTCTTTAGGTTTGACAGATGATCATATACAGAAACTAC	299
Db	252	GGGGTTGAAAGCTTGCCTTCTTTAGGTTTGACAGATGATCATATACAGAAACTAC	311
QY	300	ATCAGACAAATTTGTTGTGTGATTTCAAAATPAGAATAATAGATTAGACGGGAAAACATC	359
Db	312	ATCAGACAAATTTGTTGTGTGATTTCAAAATPAGAATAATAGATTAGACGGGAAAACATC	371
QY	360	AAAGCTTCAATA-----	371
Db	372	AAAGCTTCAATAATGGAACACAGCAGGCCAGGAAAGTTTGAACAATCATCTTCAAGTAT	431
QY	372	-----GAGTCTTCAAT	383
Db	432	TACAGAGGAGCCCATGCGATCATAGTTGTGTATGTGACACAGATCAGAGATCTTCAAT	491
QY	384	AATGTTTAAACAGTGGGTGCGAGGAAATAGATCGTTATGCGAGGAAAATCTCAACAATTG	443
Db	492	AATGTTTAAACAGTGGGTGCGAGGAAATAGATCGTTATGCGAGGAAAATCTCAACAATTG	551
QY	444	TTGTAGGGAACAAATGTGATCTGACCAACAAAGAAATGATGACTACAAACAGCAGAG	503
Db	552	TTGTAGGGAACAAATGTGATCTGACCAACAAAGAAATGATGACTACCAACAGCAGAG	611
QY	504	GAAATTGCGATTCCTTGGAAATTCGCTTTTGGAAACCAATGCTTAAGATGCAACGAAT	563
Db	612	GAAATTGCTGATTCCTTGGAAATTCGCTTTTGGAAACCAATGCTTAAGATGCAACGAAT	671
QY	564	GTAGAACAGTCTTTCATGACGATGCGAGCTGAGATTAAAGAACGAATGGGTCCGGAAGA	623
Db	672	GTAGAACAGTCTTTCATGACGATGCGAGCTGAGATTAAAGAACGAATGGGTCCGGAAGA	731
QY	624	ACAGTGGTGTGCTGAGAGTCCATGTAAAAATTCAGAGCATTCGACGTCAGAGACTCA	683
Db	732	ACAGTGGTGTGCTGAGAGTCCATGTAAAAATTCAGAGCATTCGACGTCAGAGACTCA	791
QY	684	GGTGAGAGTTCCTGCTAAATTTTGCTCCATCTTTTTCACAGCAATGAATTTGCAATC	743
Db	792	GGTGAGAGTTCCTGCTAAATTTTGCTCCATCTTTTTCACAGCAATGAATTTGCAATC	851
QY	744	TGAACCCAGTGAAGAAAACAAAATTGCTGAAATGTACTGTATGTAGCTGACTACAACA	803
Db	852	TGAACCCAGTGAAGAAAACAAAATTGCTGAAATGTACTGTATGTAGCTGACTACAACA	911
QY	804	GATTCCTTACCGTCCCAAGAGTCAAGATTTGTAAAGGTCACATACGACTTTTTTTTTT	863
Db	912	GATTCCTTACCGTCCCAAGAGTCAAGATTTGTAAAGGTCACATACGACTTTTTTTTTT	971
QY	864	ATTCCTTGACTCAAGACAGCTAATCTTCAATTTTCAGAACTGTATTTAAACCTTTGTGTCT	923
Db	972	ATTCCTTGACTCAAGACAGCTAATCTTCAATTTTCAGAACTGTATTTAAACCTTTGTGTCT	1031
QY	924	GGTTATTAATAATNGGTGTAAATCTTGTGCTTTCTGATACCAAGCTGTTTCCGCTG	983
Db	1032	GGTTATTAATAATNGGTGTAAATCTTGTGCTTTCTGATACCAAGCTGTTTCCGCTG	1091
QY	984	GTTGGTTAGATATATTTTGTTTGATGTTTATATTTGGCATGTTTAAATGTCAGGTTTAA	1043
Db	1092	GTTGGTTAGATATATTTTGTTTGATGTTTATATTTGGCATGTTTAAATGTCAGGTTTAA	1151

QY	1044	TCCTCTGAAGATGAAGGTCCAGCCATTTTGTATCAAAACGACCAAGCAGTGTCTGCACCTT	1109
QY <td>1152</td> <td>TCCTCTGAAGATGAAGGTCCAGCCATTTTGTATCAAAACGACCAAGCAGTGTCTGCACCTT</td> <td>1211</td>	1152	TCCTCTGAAGATGAAGGTCCAGCCATTTTGTATCAAAACGACCAAGCAGTGTCTGCACCTT	1211
QY <td>1104</td> <td>TCCATGCATAAAGTTTATAGTGAAGATGTATATATGTAAGATCTGATTTTCTAGTTCCTCTTG</td> <td>1163</td>	1104	TCCATGCATAAAGTTTATAGTGAAGATGTATATATGTAAGATCTGATTTTCTAGTTCCTCTTG	1163
Db <td>1212</td> <td>TCCATGCATAAAGTTTATAGTGAAGATGTATATATGTAAGATCTGATTTTCTAGTTCCTCTTG</td> <td>1271</td>	1212	TCCATGCATAAAGTTTATAGTGAAGATGTATATATGTAAGATCTGATTTTCTAGTTCCTCTTG	1271
QY <td>1164</td> <td>TAGAGTATTAATAATGGAAGAATTACATATCTGATTAATAGTTTCTTCACTACTCTGCATAT</td> <td>1223</td>	1164	TAGAGTATTAATAATGGAAGAATTACATATCTGATTAATAGTTTCTTCACTACTCTGCATAT	1223
Db <td>1272</td> <td>TAGAGTATTAATAATGGAAGAATTACATATCTGATTAATAGTTTCTTCACTACTCTGCATAT</td> <td>1331</td>	1272	TAGAGTATTAATAATGGAAGAATTACATATCTGATTAATAGTTTCTTCACTACTCTGCATAT	1331
QY <td>1224</td> <td>AAATTGTGGCTGCAGAAATATTTGTAATTTTGTTCGACACTATGTAACAAAACAACATGGAAGAT</td> <td>1283</td>	1224	AAATTGTGGCTGCAGAAATATTTGTAATTTTGTTCGACACTATGTAACAAAACAACATGGAAGAT	1283
Db <td>1332</td> <td>AAATTGTGGCTGCAGAAATATTTGTAATTTTGTTCGACACTATGTAACAAAACAACATGGAAGAT</td> <td>1391</td>	1332	AAATTGTGGCTGCAGAAATATTTGTAATTTTGTTCGACACTATGTAACAAAACAACATGGAAGAT	1391
QY <td>1284</td> <td>ATGTTTAATAAATATTGACTCTTATTTGGAGTAATAAAAAA</td> <td>1323</td>	1284	ATGTTTAATAAATATTGACTCTTATTTGGAGTAATAAAAAA	1323
Db <td>1392</td> <td>ATGTTTAATAAATATTGACTCTTATTTGGAGTAATAATAA</td> <td>1431</td>	1392	ATGTTTAATAAATATTGACTCTTATTTGGAGTAATAATAA	1431

RESULT 2	CR593530	1440 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	CR593530				
DEFINITION	full-length cDNA clone CS0DC023YH06 of Neuroblastoma Cot 25-normalized of Homo sapiens (human).				
ACCESSION	CR593530				
VERSION	CR593530.1	GI:50474337			
KEYWORDS	HTC; CNSLT_CDNA.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

REFERENCE
1 (bases 1 to 1440)
AUTHORS
Li, W. B., Gruber, C., Jesse, J., and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :

REFERENCE 2 (bases 1 to 1440)
AUTHORS Genoscope,
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Séquençage :

COMMENT

and the not 1 and back 5 sales of the pumpsforski 8 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES	Location/Qualifiers
source	1. .1440

```

/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="CS0DC023YH06"
/tissue_type="Neuroblastoma Coc 25-normalized"
/plasmid="pCMVSPORT_6"

```

Query Match	84.6%;	Score 1188;	DB 3;	Length 1440;
Best Local Similarity	93.1%;	Pred. No. 4.9e-196;		
Matches 1305; Conservative	0;	Mismatches 0;	Indels 97;	Gaps 2;

1 AAGCGATAGCTGAGT-GCGGCGGCTGCTGATTGTGTTCTAGGGGACGAGTAGGGGAAGA 59

[illegible]

Dp 99 CGTTGCTCTCCGGAACAGCCTATCTCATTCCTTCTTCGATTACCCGTCGCCGAG 158

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd

OM nucleic - nucleic search, using sw model

Run on: November 18, 2004, 18:48:15 ; Search time 749 Seconds
(without alignments)
10136.303 Million cell updates/sec

Title: US-09-820-003C-1

Sequence: 1 aagcgatagctgagtcgagc.....aaaaaaaaaaaaaaaa 1405

Scoring table: IDENTITY_NUC

Searched: 3627888 seqs, 2701811610 residues

Total number of hits satisfying chosen parameters: 7255776

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

1:	/cgn2_6/prodataa/1/pubnpa/us07_PUBCOMB.seq.*
2:	/cgn2_6/prodataa/1/pubnpa/PCST_NEW_PUB.seq.*
3:	/cgn2_6/prodataa/1/pubnpa/us06_NEW_PUB.seq.*
4:	/cgn2_6/prodataa/1/pubnpa/us06_PUBCOMB.seq.*
5:	/cgn2_6/prodataa/1/pubnpa/us07_NEW_PUB.seq.*
6:	/cgn2_6/prodataa/1/pubnpa/PCURS_PUBCOMB.seq.*
7:	/cgn2_6/prodataa/1/pubnpa/us08_NEW_PUB.seq.*
8:	/cgn2_6/prodataa/1/pubnpa/us08_PUBCOMB.seq.*
9:	/cgn2_6/prodataa/1/pubnpa/us09_PUBCOMB.seq.*
10:	/cgn2_6/prodataa/1/pubnpa/us09_PUBCOMB.seq.
11:	/cgn2_6/prodataa/1/pubnpa/us09_PUBCOMB.seq.
12:	/cgn2_6/prodataa/1/pubnpa/us09_NEW_PUB.seq.
13:	/cgn2_6/prodataa/1/pubnpa/us10_PUBCOMB.seq.
14:	/cgn2_6/prodataa/1/pubnpa/us10_PUBCOMB.seq.
15:	/cgn2_6/prodataa/1/pubnpa/us10_PUBCOMB.seq.
16:	/cgn2_6/prodataa/1/pubnpa/us10_PUBCOMB.seq.
17:	/cgn2_6/prodataa/1/pubnpa/us10_NEW_PUB.seq.
18:	/cgn2_6/prodataa/1/pubnpa/us11_NEW_PUB.seq.
19:	/cgn2_6/prodataa/1/pubnpa/us11_NEW_PUB.seq.
20:	/cgn2_6/prodataa/1/pubnpa/us60_NEW_PUB.seq.
21:	/cgn2_6/prodataa/1/pubnpa/us60_PUBCOMB.seq.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1405	100.0	1405	9	US-09-820-0034-1	Sequence 1, Appl1
2	1201.2	85.5	8131	15	US-10-252-157-411	Sequence 411, Appl
3	1103.6	78.5	2257	15	US-10-094-749-648	Sequence 648, Appl
4	994.6	70.8	1545	16	US-10-112-944-48	Sequence 48, Appl
5	817.2	58.2	46050	9	US-09-820-0034-3	Sequence 3, Appl1
6	684.4	48.7	48019	15	US-10-174-175-12	Sequence 12, Appl
7	589.8	42.0	601	9	US-09-820-0034-37	Sequence 37, Appl
8	506	36.0	506	9	US-09-920-3004-374	Sequence 374, Appl
9	506	36.0	506	13	US-10-093-528-374	Sequence 374, Appl
C 10	506	36.0	506	15	US-10-099-926-374	Sequence 374, Appl
C 11	476.2	33.9	723	16	US-10-305-720-1422	Sequence 1422, Appl
C 12	467	33.2	487	9	US-09-864-761-1742	Sequence 1742, Appl

13	412.4	29.4	414	18	US-10-425-115-167539	Sequence 167539, App
14	391.2	27.8	411	16	US-10-242-5358-35936	Sequence 35936, A
15	391.2	27.8	411	16	US-10-085-7838-35936	Sequence 35936, A
16	389.2	27.7	420	16	US-10-242-5358-47841	Sequence 47841, A
17	389.2	27.7	420	16	US-10-085-7838-47841	Sequence 47841, A
18	367	26.1	367	9	US-09-796-692-6644	Sequence 6644, App
19	367	26.1	367	14	US-10-040-862-4644	Sequence 4644, App
20	367	26.1	367	16	US-10-057-4758-4644	Sequence 4644, App
21	367	26.1	367	16	US-10-154-8848-4644	Sequence 4644, App
22	367	26.1	367	17	US-10-764-324-4644	Sequence 4644, App
23	348	24.8	360	16	US-10-242-5358-16376	Sequence 16376, A
24	348	24.8	360	16	US-10-085-7838-16376	Sequence 16376, A
25	337.8	24.0	436	10	US-09-918-995-4909	Sequence 4909, App
26	337.8	24.0	353	9	US-09-960-352-8911	Sequence 5427, App
27	333	23.7	352	16	US-10-042-5358-48372	Sequence 48372, App
28	333	23.7	352	16	US-10-085-7838-48372	Sequence 48372, A
29	331.4	23.6	357	16	US-10-242-5358-52337	Sequence 52337, A
30	331.4	23.6	357	16	US-10-085-7838-52337	Sequence 52337, A
31	314.6	22.4	427	9	US-09-960-352-8911	Sequence 8911, App
32	314.6	22.4	483	10	US-09-918-995-2704	Sequence 2704, App
33	301	21.4	1842	18	US-10-471-411-1	Sequence 1, App11
34	298.4	21.2	300	16	US-10-242-5358-58003	Sequence 58003, A
35	298.4	21.2	300	16	US-10-085-7838-58003	Sequence 58003, A
36	293	20.9	317	16	US-10-242-5358-52846	Sequence 52846, A
37	293	20.9	317	16	US-10-085-7838-52846	Sequence 52846, A
38	280.4	20.0	459	10	US-09-918-995-34631	Sequence 34631, A
39	261	18.6	290	9	US-09-294-0938-4584	Sequence 4584, App
40	260.6	18.5	297	13	US-10-079-623-231	Sequence 231, App
41	260.6	18.5	517	16	US-10-242-5358-49080	Sequence 49080, A
42	260.6	18.5	517	16	US-10-085-7838-49080	Sequence 49080, A
43	232.8	16.6	241	9	US-09-604-2878-337	Sequence 337, App
44	232.8	16.6	241	9	US-09-834-759-337	Sequence 337, App
45	232.8	16.6	241	10	US-09-551-621-337	Sequence 337, App

ALIGNMENTS

RESULT 1

; Sequence 1, Application US/09820003A

Patent No. US20020142382A1

APPLICANT: MERKULOV, Gennady et al

TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS, NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
AND NUMBER OF PAGES: 0001-0008

CURRENT APPLICATION NUMBER: US/09/820,003A

NUMBER OF SEO ID NOS: 39

```

; SOFTWARE: FA8CSEQ FOR WINDOWS VERSION 4.0
; SEQ ID NO 1

```

LENGTH: 1405

ORGANISM: Homo sapiens

US-09-820-003A-1

Query Match	100.0%;	Score 1405;	DB 9;	Length 1405;
-------------	---------	-------------	-------	--------------

```

Best Local Similarity: 100.0%; Freq: 2.1E-22;
Matches 1405: Conservative 0: Mismatches 0: Indels 0: Gaps 0

```

QY	1	AAGCGATAGCTAGTGTGCGCGGCTGTCTGATTTGTGTTCTTAGGGGACGAGATAGGGGAAAGAC	60
Db	1	AAGCGATAGCTAGTGTGCGCGGCTGTCTGATTTGTGTTCTTAGGGGACGAGATAGGGGAAAGAC	60
QY	61	GTTTGGCTTCCCGGGAACAGCTTATCTCATTTCTTTTGCATTAACCCGTGGCGCGGAGA	120
Db	61	GTTTGGCTTCCCGGGAACAGCTTATCTCATTTCTTTTGCATTAACCCGTGGCGCGGAGA	120
QY	121	GTTAAGGGCGGGCGGCTGCGGACAGCAAGGGCGGCGGCGGAGGAGGACGTCAGTGTACA	180
Db	121	GTTAAGGGCGGGCGGCTGCGGACAGCAAGGGCGGCGGCGGAGGAGGACGTCAGTGTACA	180

QY	181	TGTCAGAGATGATCCCGAATAGATTAATTTATTCAGTACTCTTGATTTGGCACTCAG	240
Db	181	TGTCAGAGATGATCCCGAATAGATTAATTTATTCAGTACTCTTGATTTGGCACTCAG	240
QY	241	GGGTTGGAAAGTCTTGCCCTTCTCTTAAGTTTGCAGATGATCATATACAGAAAGCTACA	300
Db	241	GGGTTGGAAAGTCTTGCCCTTCTCTTAAGTTTGCAGATGATCATATACAGAAAGCTACA	300
QY	301	TCAGCACAATTGGTGTGGATTTCCAAAATAAGACTATAGACTTAAACCGGAAAACAATCA	360
Db	301	TCAGCACAATTGGTGTGGATTTCCAAAATAAGACTATAGACTTAAACCGGAAAACAATCA	360
QY	361	AGCTTCAAATAGAGTCTTCAATTAATGTTAAACAGTGGCTGCAGAAATAGATCGTTATG	420
Db	361	AGCTTCAAATAGAGTCTTCAATTAATGTTAAACAGTGGCTGCAGAAATAGATCGTTATG	420
QY	421	CCAGTGGAAATGTCAACAAATTTGTTGGAGAAACAATGTGATCTGACACAAAGAAAG	480
Db	421	CCAGTGGAAATGTCAACAAATTTGTTGGAGAAACAATGTGATCTGACACAAAGAAAG	480
QY	481	TAGTAGACTACACACAGCGAAGGAATTGGATGTTCCCTGGAAATCCGTTTTGGAAA	540
Db	481	TAGTAGACTACACACAGCGAAGGAATTGGATGTTCCCTGGAAATCCGTTTTGGAAA	540
QY	541	CCAGTGTCTAAGAAATSCAAGAAATGTAAGAACAGTCTTTCATGACGATGCGAGCTGAGATTA	600
Db	541	CCAGTGTCTAAGAAATSCAAGAAATGTAAGAACAGTCTTTCATGACGATGCGAGCTGAGATTA	600
QY	601	AAAGACGAATGGGTCCCGGAGCAACAGCTGTGTGTCTGAGAAATTCATGTTAAATTC	660
Db	601	AAAGACGAATGGGTCCCGGAGCAACAGCTGTGTGTCTGAGAAATTCATGTTAAATTC	660
QY	661	AGAGCACTCCAGTCAAGCAGTCAAGTGTGCTGTCTAAATTTGCTCCATCCTTTT	720
Db	661	AGAGCACTCCAGTCAAGCAGTCAAGTGTGCTGTCTAAATTTGCTCCATCCTTTT	720
QY	721	CTCACAGCAATGGAATTTGCAATCTGAACCCCAAGTGAAGAAAACAAATTTGCCTTGAATGTGA	780
Db	721	CTCACAGCAATGGAATTTGCAATCTGAACCCCAAGTGAAGAAAACAAATTTGCCTTGAATGTGA	780
QY	781	CTGATGTAGCTGCACTCAACACAGATTCTTACCGCTCCACAAAGSTCAGAGATTGTAAA	840
Db	781	CTGATGTAGCTGCACTCAACACAGATTCTTACCGCTCCACAAAGSTCAGAGATTGTAAA	840
QY	841	TGTCGAATACTGACTTTTATTTTATTTCCCTTGACTCAAGACGCTTAACTTCATTTTCAGA	900
Db	841	TGTCGAATACTGACTTTTATTTTATTTCCCTTGACTCAAGACGCTTAACTTCATTTTCAGA	900
QY	901	ACTGTTTAAACCTTTGTGTGTGCTGTTATATTAATATGTTGTATATCTTGTTGCTTTC	960
Db	901	ACTGTTTAAACCTTTGTGTGTGCTGTTATATTAATATGTTGTATATCTTGTTGCTTTC	960
QY	961	CTGATACCAAGCTGTTCGCCGTGTGGTTAAGATTAATTTTGTTTGATGTTTATATTTG	1020
Db	961	CTGATACCAAGCTGTTCGCCGTGTGGTTAAGATTAATTTTGTTTGATGTTTATATTTG	1020
QY	1021	GCATGTTTAAGATGCAGAGTTTAAGCTTCTCTGAAGATGAAGTTCACGCCATTTTGTATCAAC	1080
Db	1021	GCATGTTTAAGATGCAGAGTTTAAGCTTCTCTGAAGATGAAGTTCACGCCATTTTGTATCAAC	1080
QY	1081	AGCACAAGCAGTGTCTGTCACTTCCCAAGTCAATTAAGTTTGTGAGATTAATATGTATAGA	1140
Db	1081	AGCACAAGCAGTGTCTGTCACTTCCCAAGTCAATTAAGTTTGTGAGATTAATATGTATAGA	1140
QY	1141	TCGTGATTTGCTAGTCTTCCCTGTGAGATTAAATGAAAGATTACACTATCTGATTTAA	1200
Db	1141	TCGTGATTTGCTAGTCTTCCCTGTGAGATTAAATGAAAGATTACACTATCTGATTTAA	1200
QY	1201	TAGTTTCTTATCTCGCATATTAATTTGTGGCTGCAGAATATTGTATATTTGTTGCACAC	1260
Db	1201	TAGTTTCTTATCTCGCATATTAATTTGTGGCTGCAGAATATTGTATATTTGTTGCACAC	1260

Qy	1261	TATGTACAAAACA	CTGAGATAGCTTAAATAATATGACTATATGGAAGTAA	1320
Db	1261	TATGTACAAAACA	CTGAGATAGCTTAAATAATATGACTATATGGAAGTAA	1320
Qy	1321	AAAAAAAAAAAAA	AA	1360
Db	1321	AAAAAAAAAAAAA	AA	1360
Qy	1381	AAAAAAAAAAAAA	AA	1405
Db	1381	AAAAAAAAAAAAA	AA	1405
RESULT 2				
US-10-252-157-411				
/ Sequence 411, Application US/10252157				
/ Publication No. US20030190640A1				
/ GENERAL INFORMATION:				
/ APPLICANT: Paris, Mary				
/ APPLICANT: Pearson, Cecelia I.				
/ TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER				
/ FILE REFERENCE: PA-0027-1 US				
/ CURRENT APPLICATION NUMBER: US/10/252,157				
/ PRIOR FILING DATE: 2002-10-01				
/ PRIOR APPLICATION NUMBER: 60/295,048				
/ PRIOR FILING DATE: 2001-05-31				
/ NUMBER OF SEQ ID NOS: 501				
/ SOFTWARE: PERL Program				
/ SEQ ID NO 411				
/ LENGTH: 8131				
/ TYPE: DNA				
/ ORGANISM: Homo sapiens				
/ FEATURE:				
/ NAME/KEY: misc.feature				
/ OTHER INFORMATION: Incycle ID No. US20030190640A1 411474.17				
/ FEATURE:				
/ NAME/KEY: unsure				
/ LOCATION: 3486-3786				
/ OTHER INFORMATION: a, t, c, g, or other				
US-10-252-157-411				
Query Match 85.5%; Score 1201.2; DB 15; Length 8131;				
Best Local Similarity 93.0%; Pred. No. 9e-193; Indels 97; Gaps 2;				
Matches 1320; Conservative 0; Mismatches 3;				
Qy	1	AAGGATAGCTGAGT	-GGCGCGCTGCTGATTTGTCTTAAAGGACGAGTAGGGAGA	59
Db	21	AAGGATAGCTGAGT	GGCGCGCTGCTGATTTGTCTTAAAGGACGAGTAGGGAGA	80
Qy	60	CGTTTGCTCCCGGAA	CACCTATCTCATTTCTTTCTTTCGATTCACCGTGGCGGGAG	119
Db	81	CGTTTGCTCCCGGAA	CAAGCCTATCTCATTTCTTTCTTTCGATTCACCGTGGCGGGAG	140
Qy	120	AGTCAGGCGCGCG	CGCTGCGGACAAAGGCGCGGTGGCGCGCGCGCGAGCTGCATGAC	179
Db	141	AGTCAGGCGCGCG	CGCTGCGGACAAAGGCGCGGTGGCGCGCGCGCGAGCTGCATGAC	200
Qy	180	ATGTCGACGATGAAT	CCCGAATATGATTTTATTTCAATTAATCTTCTGATTTGGCACTCA	239
Db	201	ATGTCGACGATGAAT	CCCGAATATGATTTTATTTCAATTAATCTTCTGATTTGGCACTCA	260
Qy	240	GGGGTTGGAAAGT	CTTGCTTTCTTTAGGTTTGCAGATATACATATACAGAAAGCTAC	299
Db	261	GGGGTTGGAAAGT	CTTGCTTTCTTTAGGTTTGCAGATATACATATACAGAAAGCTAC	320
Qy	300	ATAGACAACTATG	TGTGATTTCAAAAATAAGACTATATGAGTTGACGGGAAAACAATC	359
Db	321	ATAGACAACTATG	TGTGATTTCAAAAATAAGACTATATGAGTTGACGGGAAAACAATC	380
Qy	360	AAAGCTTCAATA	-----	371
Db	381	AAAGCTTCAATA	-----	440

DR P-PSDB; ABG72742.

XX Novel isolated human Ras-like polypeptide useful for diagnosing,
PT preventing and treating inflammation and disorders associated with cell
PT proliferation and apoptosis.

XX Claim 4; Fig 1; 70pp; English.

XX The invention discloses an isolated human Ras-like polypeptide and the
XX polynucleotide that encodes it. Ras proteins are small (low molecular
XX weight, LMW) regulatory guanine nucleotide-binding protein (GTP-binding
XX proteins or small G proteins) and are key relays in the signal
XX transduction cascade. The Ras-like proteins participate in a wide range
XX of regulatory functions. The Ras-like protein is useful for identifying a
XX modulator of function or activity and for identifying an agent that binds
XX to it. The polypeptide and antibody are useful in the diagnosis,
XX prevention and treatment of disorders associated with an increase in
XX apoptosis, such as AIDS, neurodegenerative diseases, such as Alzheimer's
XX disease, Parkinson's disease, myelodysplastic syndromes, such as aplastic
XX anaemia, ischaemic injuries, such as myocardial infarction, stroke, toxin
XX -induced diseases, such as alcohol-induced liver damage, cirrhosis,
XX and C and osteoporosis, cell proliferation disorders, such as hepatitis B
XX leukemia and inflammation disorders, such as allergy, asthma,
XX atherosclerosis, diabetes mellitus and rheumatoid arthritis. The
XX polypeptide and polynucleotide are useful as models for the development
XX of human therapeutics. The polypeptide is useful to raise antibodies or
XX biological fluids and tissues, in drug screening assays and in
XX pharmacogenomic analysis. The polynucleotide is useful for constructing
XX transgenic animals, for monitoring the effectiveness of modulating
XX compounds, for testing an individual for a genotype, as antisense
XX constructs and for gene therapy. The sequence presented is the human Ras-
XX like protein cDNA

XX Sequence 1405 BP; 471 A; 234 C; 300 G; 400 T; 0 U; 0 Other;

Query Match 100.0%; Score 1405; DB 10; Length 1405;
Best Local Similarity 100.0%; Pred. No. 2e-183;
Matches 1405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AAGCATAGCTGAGTGGCGGCGCTGATGTTGTTCAAGGGGAGAGTGAAGGAAAC 60
DB 1 AAGCATAGCTGAGTGGCGGCGCTGATGTTGTTCAAGGGGAGAGTGAAGGAAAC 60
QY 61 GTTGTCTCTCCCGAAGACGCTATCTCATCTCTTCTTGAATTAACCGTGGCGGAGA 120
DB 61 GTTGTCTCTCCCGAAGACGCTATCTCATCTCTTCTTGAATTAACCGTGGCGGAGA 120
QY 121 GTCAGGCGGCGGCGCTGGCGGAGCAAGGGCGGCGTGGCGGCGGAGCTGCAATGACA 180
DB 121 GTCAGGCGGCGGCGCTGGCGGAGCAAGGGCGGCGTGGCGGCGGAGCTGCAATGACA 180
QY 181 TGTTCAGCATGATCCGGAATATGATTAATTAATTAATTAATTAATTAATTAATTA 240
DB 181 TGTTCAGCATGATCCGGAATATGATTAATTAATTAATTAATTAATTAATTAATTA 240
QY 241 GGGTTGGAAGAGTCTTGGCTCTCTTAAGTTTGAAGTATACATATACAGAAAGTACA 300
DB 241 GGGTTGGAAGAGTCTTGGCTCTCTTAAGTTTGAAGTATACATATACAGAAAGTACA 300
QY 301 TCAGCACAATTTGGTGGATTTCAAAATAGAACTATAGTTTGAAGGAGAAACATCA 360
DB 301 TCAGCACAATTTGGTGGATTTCAAAATAGAACTATAGTTTGAAGGAGAAACATCA 360
QY 361 AGCTTCAAAATAGAGTCTTCAATTAATGTTTAAACAGTGGCTGACAGAAATAGATCGTTATG 420
DB 361 AGCTTCAAAATAGAGTCTTCAATTAATGTTTAAACAGTGGCTGACAGAAATAGATCGTTATG 420
QY 421 CCAGTGAATAATGTCACAAATTTGTTGGTGAAGCAAAATGATATGACCAAAAGAAAG 480
DB 421 CCAGTGAATAATGTCACAAATTTGTTGGTGAAGCAAAATGATATGACCAAAAGAAAG 480

```

```

QY 481 TAGTAGACTACACACAGCGAAGGAAATTTGCTGATTCCTTGGAAATTCGTTTGGAAA 540
DB 481 TAGTAGACTACACACAGCGAAGGAAATTTGCTGATTCCTTGGAAATTCGTTTGGAAA 540
QY 541 CCAGTGTCAAGATGACACAGATGTAGAAACAGTCTTTCAAGAGATGGGACGTGAGATTA 600
DB 541 CCAGTGTCAAGATGACACAGATGTAGAAACAGTCTTTCAAGAGATGGGACGTGAGATTA 600
QY 601 AAAGCGAATGGGTCCCGGAGCAACAGCTGTGTGCTGAGAAATGTTAAATTC 660
DB 601 AAAGCGAATGGGTCCCGGAGCAACAGCTGTGTGCTGAGAAATGTTAAATTC 660
QY 661 AGAGCACTCAGTCAAGCAAGTCAAGTGGAGGTTGCTGTAAATTTGCTTCATCTTTT 720
DB 661 AGAGCACTCAGTCAAGCAAGTCAAGTGGAGGTTGCTGTAAATTTGCTTCATCTTTT 720
QY 721 CTACAGCAATGAATTTGCAATCTGAACCCAAAGTAAAAACAAATTTGCTGAATTTGA 780
DB 721 CTACAGCAATGAATTTGCAATCTGAACCCAAAGTAAAAACAAATTTGCTGAATTTGA 780
QY 781 CTGTATGTAGCTGCACTACACAGATTTTACCGTCTCCACAAAGTCAAGATTTGAAA 840
DB 781 CTGTATGTAGCTGCACTACACAGATTTTACCGTCTCCACAAAGTCAAGATTTGAAA 840
QY 841 TGGTCAATCTGACTTTTATTTTATTTCCCTTGACTCAAGACAGTAACTTCAATTTTGA 900
DB 841 TGGTCAATCTGACTTTTATTTTATTTCCCTTGACTCAAGACAGTAACTTCAATTTTGA 900
QY 901 ACTGTTTAAACCTTTGCTGTGCTGTTTAAATTAATGTTGTTATCTTGTGCTTTC 960
DB 901 ACTGTTTAAACCTTTGCTGTGCTGTTTAAATTAATGTTGTTATCTTGTGCTTTC 960
QY 961 CTGATACCAAGTCTTTCCCGTGGTGGTGAATAATTTTGTGTTGATGTTATATG 1020
DB 961 CTGATACCAAGTCTTTCCCGTGGTGGTGAATAATTTTGTGTTGATGTTATATG 1020
QY 1021 GGATGTTTGAAGTCAAGTGTAAATTTAGTCTTGAAGATGAAGTTCAAGCAATTTGTAAC 1080
DB 1021 GGATGTTTGAAGTCAAGTGTAAATTTAGTCTTGAAGATGAAGTTCAAGCAATTTGTAAC 1080
QY 1081 AGCACAACAGTGTCTGCTCACTTCCAGTCAATGAAGTTTATGATGATTTATGATGA 1140
DB 1081 AGCACAACAGTGTCTGCTCACTTCCAGTCAATGAAGTTTATGATGATTTATGATGA 1140
QY 1141 TCTGATTTGCTAGTCTTCTCTTGTGAGTATTAATGAAGATTAACATATCTGATTTA 1200
DB 1141 TCTGATTTGCTAGTCTTCTCTTGTGAGTATTAATGAAGATTAACATATCTGATTTA 1200
QY 1201 TAGTTTCTTCAATCTGCAATTAATTTGGCTGCAAGATTTGTTGTTGTTGTTGTTGTTG 1260
DB 1201 TAGTTTCTTCAATCTGCAATTAATTTGGCTGCAAGATTTGTTGTTGTTGTTGTTGTTG 1260
QY 1261 TATGTAAACAAACAGTGAAGATATGTTTAAATTAATTTGTTGTTGTTGTTGTTGTTG 1320
DB 1261 TATGTAAACAAACAGTGAAGATATGTTTAAATTAATTTGTTGTTGTTGTTGTTGTTG 1320
QY 1321 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1380
DB 1321 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1380
QY 1381 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1405
DB 1381 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1405

```

RESULT 2
ADP03021 standard; cDNA; 1433 BP.
ADP03021;
29-JUL-2004 (first entry)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2004, 09:13:54 ; Search time 6027 Seconds
(without alignments)
11024.063 Million cell updates/sec

Title: US-09-820-003C-1

Perfect score: 1405

Sequence: 1 aagcgatagctgagtcgagcgc.....aaaaaaaaaaaaaaaaaaaaa 1405

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*
1: gb_ha:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1277	90.9	1625	9 HSM806623	BX571747 Homo sapi
2	1219	86.8	1433	9 BC000905	BC000905 Homo sapi
3	1179.2	83.9	8137	6 AR454534	AR454534 Sequence
4	1103.6	78.5	2257	6 AX713964	AX713964 Sequence
5	1103.6	78.5	2257	9 AK055927	AK055927 Homo sapi
6	1052.6	74.9	1444	10 BC002077	BC002077 Mus muscu
7	1044	74.3	1208	9 HSM800023	AL050268 Homo sapi
8	1036.2	73.8	2819	10 AK129477	AK129477 Mus muscu
9	1028.2	73.2	1425	10 BC066662	BC066662 Rattus no
10	1013.2	72.1	1428	10 MMYPT1	Y00094 Mouse mRna
11	955	68.0	2686	10 AF226873	AF226873 Mus muscu
12	820.6	58.4	293895	2 AC123143	AC123143 Rattus no
13	817.2	58.2	204230	2 AC007318	AC007318 Homo sapi
14	797.4	56.8	329753	2 AC096701	AC096701 Rattus no
15	777.8	55.4	327906	2 AC117841	AC117841 Rattus no
16	741.4	52.8	203071	10 AL606522	AL606522 Mouse DNA
17	727.8	51.8	243263	2 AC132510	AC132510 Rattus no
18	697.2	49.6	327906	2 AC117841	AC117841 Rattus no
19	687.6	48.9	116614	2 AL512349	AL512349 Homo sapi

c 20	686	48.8	104081	9 HSDJ2637	AL049545 Human DNA
c 21	684.4	48.7	162471	9 AC067945	AC067945 Homo sapi
c 22	682.6	48.6	4678	10 MMYPT14	X15747 Mouse YP1
c 23	674	48.0	217009	2 AC094972	AC094972 Rattus no
c 24	651	46.3	703	11 BV166639	BV166639 RBIA_104
c 25	622.6	44.3	1006	11 BV105532	BV105532 MARC_2615
c 26	611.6	43.5	191930	2 AC120240	AC120240 Rattus no
c 27	573.8	40.8	310641	2 AC099354	AC099354 Rattus no
c 28	573.2	40.8	840	4 CFRAB1	X56384 Canine rab1
c 29	567.6	40.4	141036	2 AC023807	AC023807 Mus muscu
c 30	543.4	38.7	1086	10 RN063023	U63023 Rattus norv
c 31	506	36.0	506	6 AX396159	AX396159 Sequence
c 32	476.2	33.9	723	6 AR270859	AR270859 Sequence
c 33	476.2	33.9	723	9 HUMRAB1A	M28209 Homo saplen
c 34	467	33.2	487	6 CQ050923	CQ050923 Sequence
c 35	467	33.2	487	6 CQ065951	CQ065951 Sequence
c 36	467	33.2	487	6 CQ092963	CQ092963 Sequence
c 37	467	33.2	487	6 CQ131774	CQ131774 Sequence
c 38	467	33.2	487	6 CQ170346	CQ170346 Sequence
c 39	467	33.2	487	6 CQ199514	CQ199514 Sequence
c 40	467	33.2	487	6 CQ214966	CQ214966 Sequence
c 41	467	33.2	487	6 CQ253556	CQ253556 Sequence
c 42	467	33.2	487	6 CQ290658	CQ290658 Sequence
c 43	467	33.2	487	6 CQ327680	CQ327680 Sequence
c 44	455.8	32.4	110000	2 AC096343	AC096343 Rattus no
c 45	452.2	32.2	110000	2 AC094411_1	Continuation (2 of

ALIGNMENTS

RESULT 1
HSM806623
LOCUS
DEFINITION
Homo sapiens mRNA; cDNA DKFZp686E08159 (from clone DKFZp686E08159); complete cds.
ACCESSION
BX571747
KEYWORDS
BX571747.1 GI:33096734
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1625)
Koehler,K., Beyer,A., Mewes,H.W., Weil,B., Auld,C., Osanger,A., Fodor,G., Han,W. and Wiemann,S.
The German Human cDNA Consortium
Direct Submission
Submitted (16-JUL-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Heinrich-Heine-University, Dueseldorf(Germany) within the CDNA sequencing consortium of the German Genome Project. This clone (DKFZp686E08159) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heudneweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cdna/>.
FEATURES
source
1..1625
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="2p14"
/clone="DKFZp686E08159"
/tissue_type="human colon endothel primary cell culture"
/clone_lib="686 (synonym: hlcc3). Vector poport1_sfi; host DH10B; sites SfiI + SfiIb"
/dev_stage="adult"
1..1625
/gene="DKFZp686E08159"

CDS

```

184..801
/gene="DKFZp686E08159"
/locus="RAB1A, member RAS oncogene family"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAE11872.1"
/db_xref="GI:3306735"
/translation="MSSMNPEDYLFKLLIGDSGVKSCILRFADDTYESYIT
KMDLPIRINSEVNNKLLVGNKCDLTKKVDYDTAKFEADSLGIPLETSAKMN
VEQSPMTAAELKRMGPGATAGAKSNVXIQSTPVKQSGGGCC"
polya_signal
1391..1396
/gene="DKFZp686E08159"
polya_site
1411
/gene="DKFZp686E08159"
ORIGIN

```

```

Query Match      90.9%; Score 1277; DB 9; Length 1625;
Best Local Similarity 93.5%; Pred. No. 1e-194;
Matches 1405; Conservative 0; Mismatches 0; Indels 98; Gaps 3;

```

```

1  AAGGATAGCTGAGT-GGGGCGCTGCTGATTTGTTCTTGAGGAGCGAGTAGGGAGAGA 59
Db  4  AAGGATAGCTGAGTGGCGGCTGCTGATTTGTTCTTGAGGAGCGAGTAGGGAGAGA 63
Qy  60  CGTTGCTCTCCGGAAACAGCTATCTCATTTCTTTCGATTACCGGTGGCGCGAG 119
Db  64  CGTTGCTCTCCGGAAACAGCTATCTCATTTCTTTCGATTACCGGTGGCGCGAG 123
Qy  120  AGTCAGGCGGCGGCTGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 179
Db  124  AGTCAGGCGGCGGCTGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 183
Qy  180  ATGTCCAGCTGATCCGAATATGATTTATTTCAAGTTACTCTGATGGGCACTCA 239
Db  184  ATGTCCAGCTGATCCGAATATGATTTATTTCAAGTTACTCTGATGGGCACTCA 243
Qy  240  GGGGTTGAAAAGCTCTCTCTCTGAGTTTGCAGATGATATACATATACAGAAAGCTAC 299
Db  244  GGGGTTGAAAAGCTCTCTCTCTGAGTTTGCAGATGATATACATATACAGAAAGCTAC 303
Qy  300  ATCAGACATTTGGTGTGATTTCAAAATAGAACTATAGAGTTAGACGGGAAACATC 359
Db  304  ATCAGACATTTGGTGTGATTTCAAAATAGAACTATAGAGTTAGACGGGAAACATC 363
Qy  360  AAGCTTCAAAATA----- 371
Db  364  AAGCTTCAAAATATGGGACACAGCGCCAGGAAAGATTGGAACAATCACCTCAAGTTAT 423
Qy  372  -----GAGTCCCTTCAAT 383
Db  424  TACAGAGAGCCCATGCGATCATAGTTGTGATGATGTGACAGATCAGGAGTCTTCAAT 483
Qy  384  AATGTTAAACAGTGGCTGCGAGGAAATAGATCGTTATGCCAGTAAATGTCACAAATTG 443
Db  484  AATGTTAAACAGTGGCTGCGAGGAAATAGATCGTTATGCCAGTAAATGTCACAAATTG 543
Qy  444  TTGCTAGGGAACAATATGATCTGACCAAGAAAGTATGATACACACACAGAGAG 503
Db  544  TTGCTAGGGAACAATATGATCTGACCAAGAAAGTATGATACACACACAGAGAG 603
Qy  504  GAATTTGCTGATTCCTTGGAAATTCGTTTGGAAACAGTGTCAAGATGCAACGAT 563
Db  604  GAATTTGCTGATTCCTTGGAAATTCGTTTGGAAACAGTGTCAAGATGCAACGAT 663
Qy  564  GTTGAACAGTCTTTCATGACGATGCGAGCTGAGATTTAAAGCGATGGTCCCGAGCA 723
Db  664  GTTGAACAGTCTTTCATGACGATGCGAGCTGAGATTTAAAGCGATGGTCCCGAGCA 723
Qy  624  ACAGCTGTGTGCTGAGAGTTCATGTTAAATTCAGAGACTTCAGTCAAGAGTCA 683
Db  724  ACAGCTGTGTGCTGAGAGTTCATGTTAAATTCAGAGACTTCAGTCAAGAGTCA 783

```

```

Qy  684  GGTGAGAGTTGCTGTAAATTTTGCTCCATCTTTCTACAGACATGAAATTTGCAATC 743
Db  784  GGTGAGAGTTGCTGTAAATTTTGCTCCATCTTTCTACAGACATGAAATTTGCAATC 843
Qy  744  TGAACCAAGTGAAGAAACAAATTTGCTGAAATTTGATCTGATGAGTCTGACATCAACA 803
Db  844  TGAACCAAGTGAAGAAACAAATTTGCTGAAATTTGATCTGATGAGTCTGACATCAACA 903
Qy  804  GATTTTACCGCTCTCCACAAAGGTCAGAGATTTGAAATGTCATATCTGAC-TTTTTTTT 862
Db  904  GATTTTACCGCTCTCCACAAAGGTCAGAGATTTGAAATGTCATATCTGAC-TTTTTTTT 963
Qy  863  TATTCCTGAGCTCAACACGCTTACTTCAATTTTCAGAACTGTTTAAACCTTTGTCG 922
Db  964  TATTCCTGAGCTCAACACGCTTACTTCAATTTTCAGAACTGTTTAAACCTTTGTCG 1023
Qy  923  TGTGTTTAAATATATGTTGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 982
Db  1024  TGTGTTTAAATATATGTTGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1083
Qy  983  GGTGTTTAAATATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1042
Db  1084  GGTGTTTAAATATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1143
Qy  1043  GTCTTGGAAGATGAAAGTTCAGCAATTTTGTATCAACAGCAACAGCGTCTGTCACT 1102
Db  1144  GTCTTGGAAGATGAAAGTTCAGCAATTTTGTATCAACAGCAACAGCGTCTGTCACT 1203
Qy  1103  TTCCATGCAATTAAGTTAGTGAATGTTATATGTAAGATCTGATTTGCTGTTCTCTT 1162
Db  1204  TTCCATGCAATTAAGTTAGTGAATGTTATATGTAAGATCTGATTTGCTGTTCTCTT 1263
Qy  1163  GTAGAGTTTAAATGGAAGATTCACATCTGATTAATGTTTCTTCATCTCTGCATA 1222
Db  1264  GTAGAGTTTAAATGGAAGATTCACATCTGATTAATGTTTCTTCATCTCTGCATA 1323
Qy  1223  TAAATTTGGCTGCAATATTTGTAATTTGTCACATCTGTAACAAACATGTAAG 1282
Db  1324  TAAATTTGGCTGCAATATTTGTAATTTGTCACATCTGTAACAAACATGTAAG 1383
Qy  1283  TATGTTTAAATATATGTTACTTATTTGAGTAAATTAATGTTTCTTCATCTCTGCATA 1342
Db  1384  TATGTTTAAATATATGTTACTTATTTGAGTAAATTAATGTTTCTTCATCTCTGCATA 1443
Qy  1343  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1402
Db  1444  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1503
Qy  1403  AAA 1405
Db  1504  AAA 1506

```

RESULT 2
 LOCUS BC000905 1433 bp mRNA linear PRI 29-UN-2004
 DEFINITION Homo sapiens RAB1A, member RAS oncogene family, mRNA (cDNA clone
 ACCESION BC000905 IMAGE:290705), complete cds.
 VERSION BC000905.2 GI:14705268
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 1 (bases 1 to 1433)
 AUTHORS Klausner,R.D., Collins,F.S., Wagner,L.H., Derge,D.G.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diachenko,L., Maruine,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Ueda,T.B., Toshiyuki,S.,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 14:17:35 ; Search time 140 Seconds
(without alignments)
437.601 Million cell updates/sec

Title: US-09-820-003C-2

Perfect score: 886

Sequence: 1 MSSNNPEYDYLFKLLIGDS.....EKSNNVXIQSTPVKQSGG6CC 173

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBSCOMB.pep:*

2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubppaa/US06_PUBSCOMB.pep:*

5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBSCOMB.pep:*

7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubppaa/US08_PUBSCOMB.pep:*

9: /cgn2_6/ptodata/2/pubppaa/US09_PUBSCOMB.pep:*

10: /cgn2_6/ptodata/2/pubppaa/US09_PUBSCOMB.pep:*

11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBSCOMB.pep:*

12: /cgn2_6/ptodata/2/pubppaa/US09C_PUBSCOMB.pep:*

13: /cgn2_6/ptodata/2/pubppaa/US10_PUBSCOMB.pep:*

14: /cgn2_6/ptodata/2/pubppaa/US10_PUBSCOMB.pep:*

15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBSCOMB.pep:*

16: /cgn2_6/ptodata/2/pubppaa/US10D_PUBSCOMB.pep:*

17: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*

18: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep:*

19: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*

20: /cgn2_6/ptodata/2/pubppaa/US60_PUBSCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	886	100.0	173	9	US-09-820-003A-2
2	886	100.0	173	15	US-10-112-944-292
3	860	97.1	222	9	US-09-820-003A-4
4	857	96.7	205	16	US-10-804-491-49
5	812.5	91.7	201	17	US-10-471-411-6
6	792.5	89.4	169	17	US-10-471-411-2
7	768.5	86.7	224	9	US-09-925-298-466
8	768.5	86.7	224	14	US-10-102-806-466
9	766.5	86.5	201	9	US-09-967-736-3
10	766.5	86.5	201	14	US-10-291-172-193
11	766.5	86.5	201	15	US-10-221-278-193
12	766.5	86.5	201	16	US-10-408-765A-1596
13	766.5	86.5	201	17	US-10-471-411-4

14	760.5	85.8	201	9	US-09-967-736-8	Sequence 8, Appl1
15	759.5	85.7	201	17	US-10-471-411-5	Sequence 5, Appl1
16	700	79.0	141	17	US-10-094-749-2287	Sequence 2287, Appl1
17	678	76.5	204	17	US-10-471-411-7	Sequence 7, Appl1
18	660	74.5	203	14	US-10-369-493-6261	Sequence 6261, Appl1
19	658.5	74.3	203	17	US-10-471-411-8	Sequence 8, Appl1
20	645	72.8	202	15	US-10-424-599-196551	Sequence 196551, Appl1
21	635	71.7	225	15	US-10-128-714-8241	Sequence 8241, Appl1
22	624	70.4	225	15	US-10-425-114-47782	Sequence 47782, Appl1
23	624	70.4	239	15	US-10-425-114-41359	Sequence 41359, Appl1
24	624	70.4	240	15	US-10-425-114-42447	Sequence 42447, Appl1
25	624	70.4	240	15	US-10-425-114-56916	Sequence 56916, Appl1
26	623	70.3	202	15	US-10-424-599-186767	Sequence 186767, Appl1
27	622	70.2	203	17	US-10-425-115-362779	Sequence 362779, Appl1
28	621.5	70.1	203	16	US-10-767-701-445295	Sequence 445295, Appl1
29	621.5	70.1	203	17	US-10-425-115-318489	Sequence 318489, Appl1
30	621.5	70.1	225	15	US-10-425-114-47497	Sequence 47497, Appl1
31	621.5	70.1	225	15	US-10-425-114-60367	Sequence 60367, Appl1
32	621.5	70.1	225	15	US-10-425-114-67150	Sequence 67150, Appl1
33	620.5	70.0	213	15	US-10-424-599-186764	Sequence 186764, Appl1
34	617	69.6	203	16	US-10-767-701-44811	Sequence 44811, Appl1
35	616.5	69.6	202	16	US-10-437-963-123354	Sequence 123354, Appl1
36	615.5	69.5	203	17	US-10-425-115-318493	Sequence 318493, Appl1
37	615.5	69.5	228	15	US-10-425-114-60163	Sequence 60163, Appl1
38	615.5	69.5	228	15	US-10-425-114-62895	Sequence 62895, Appl1
39	607.5	68.6	205	15	US-10-425-114-61060	Sequence 61060, Appl1
40	605	68.6	204	17	US-10-424-599-186768	Sequence 186768, Appl1
41	599	67.6	216	15	US-10-424-599-186768	Sequence 186768, Appl1
42	596.5	67.3	207	14	US-10-032-585-7303	Sequence 7303, Appl1
43	591.5	66.8	212	17	US-10-425-115-200936	Sequence 200936, Appl1
44	589	66.5	304	16	US-10-437-963-136472	Sequence 136472, Appl1
45	578.5	65.3	205	17	US-10-425-115-331973	Sequence 331973, Appl1

ALIGNMENTS

RESULT 1

US-09-820-003A-2

Sequence 2, Application US/09820003A

Patent No. US20020142382A1

GENERAL INFORMATION:

APPLICANT: MERCK/COV, Genmady et al.

TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

FILE REFERENCE: CLO01196

CURRENT APPLICATION NUMBER: US/09/820,003A

CURRENT FILING DATE: 2001-03-29

NUMBER OF SEQ ID NOS: 39

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 2

LENGTH: 173

TYPE: PRT

ORGANISM: Homo sapien

US-09-820-003A-2

Query Match	100.0%;	Score 886;	DB 9;	Length 173;
Best local similarity	100.0%;	Pred. No. 3.9e-78;		
Matches 173;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MSSNNPEYDYLFKLLIGDSGVGSCCLLPADDTYESTYSTTGVDFFKIRTELDGKTI	60	
DB	1	MSSNNPEYDYLFKLLIGDSGVGSCCLLPADDTYESTYSTTGVDFFKIRTELDGKTI	60	
QY	61	KLQIESNNYKQWLOEIDRYASEVNNKLAVGNKCDLTTKKVVDYTTAKEFADSLGIPFLE	120	
DB	61	KLQIESNNYKQWLOEIDRYASEVNNKLAVGNKCDLTTKKVVDYTTAKEFADSLGIPFLE	120	
QY	121	TSANATNVEQSFWTMAEIKKMGPGATAGAEKSNVKIQSTPVKQSGG6CC	173	
DB	121	TSANATNVEQSFWTMAEIKKMGPGATAGAEKSNVKIQSTPVKQSGG6CC	173	

```
RESULT 2
US-10-112-944-292
; Sequence 292, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Meng, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1e1 Nucleic Acids and
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: Pf_Fl_genes Version 5.0
; SEQ ID NO 292
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-112-944-292

Query Match      100.0%; Score 886; DB 15; Length 173;
Best Local Similarity 100.0%; Pred. No. 3,9e-78;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSNNPEYDYLFKLLIGDSGVGSKCLLRFPADDTYTESYSTIGVDPKIRITELDGKTI 60
DB 1 MSSNNPEYDYLFKLLIGDSGVGSKCLLRFPADDTYTESYSTIGVDPKIRITELDGKTI 60
QY 61 KLOIESNNYKQWLOEIDRYASENVNKLIVGNKCDLTTKKVVDYTTAKEFADSLGIPIELE 120
DB 61 KLOIESNNYKQWLOEIDRYASENVNKLIVGNKCDLTTKKVVDYTTAKEFADSLGIPIELE 120
QY 121 TSAKNATNVBOSFMTMAAEIKKMGPGATAGAEKSNVKTOSTPVKQSGGGCC 173
DB 121 TSAKNATNVBOSFMTMAAEIKKMGPGATAGAEKSNVKTOSTPVKQSGGGCC 173

RESULT 3
US-09-820-003A-4
; Sequence 4, Application US/09820003A
; Patent No. US20020142382A1
; GENERAL INFORMATION:
; APPLICANT: MERRILOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
```

```
; FILE REFERENCE: CL001196
; CURRENT APPLICATION NUMBER: US/09/820,003A
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-820-003A-4

Query Match      97.1%; Score 860; DB 9; Length 222;
Best Local Similarity 84.4%; Pred. No. 1,8e-75;
Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSNNPEYDYLFKLLIGDSGVGSKCLLRFPADDTYTESYSTIGVDPKIRITELDGKTI 60
DB 18 MSSNNPEYDYLFKLLIGDSGVGSKCLLRFPADDTYTESYSTIGVDPKIRITELDGKTI 77
QY 61 KLOI-----ESFNNVQWLOEIDRYASENVNKL 88
DB 78 KLOIWDAGQERFRITSSYRGANGIIVYVDYDQESFNNVQWLOEIDRYASENVNKL 137
QY 89 LVGNKCDLTTKKVVDYTTAKEFADSLGIPIELTSAKNATNVBOSFMTMAAEIKKMGPGA 148
DB 138 LVGNKCDLTTKKVVDYTTAKEFADSLGIPIELTSAKNATNVBOSFMTMAAEIKKMGPGA 197
QY 149 TAGAEKSNVKTOSTPVKQSGGGCC 173
DB 198 TAGAEKSNVKTOSTPVKQSGGGCC 222

RESULT 4
US-10-804-491-49
; Sequence 49, Application US/10804491
; Publication No. US20040180375A1
; GENERAL INFORMATION:
; APPLICANT: Cismowski, Mary
; TITLE OF INVENTION: AGS Proteins and Nucleic Acid Molecules and Uses Therefor
; FILE REFERENCE: 60388-A-PCT-US
; CURRENT APPLICATION NUMBER: US/10/804,491
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: US/09/709,103
; PRIOR FILING DATE: 2000-11-08
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-804-491-49

Query Match      96.7%; Score 857; DB 16; Length 205;
Best Local Similarity 83.9%; Pred. No. 3,3e-75;
Matches 172; Conservative 1; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSNNPEYDYLFKLLIGDSGVGSKCLLRFPADDTYTESYSTIGVDPKIRITELDGKTI 60
DB 1 MSSNNPEYDYLFKLLIGDSGVGSKCLLRFPADDTYTESYSTIGVDPKIRITELDGKTI 60
QY 61 KLOI-----ESFNNVQWLOEIDRYASENVNKL 88
DB 61 KLOIWDAGQERFRITSSYRGANGIIVYVDYDQESFNNVQWLOEIDRYASENVNKL 120
QY 89 LVGNKCDLTTKKVVDYTTAKEFADSLGIPIELTSAKNATNVBOSFMTMAAEIKKMGPGA 148
DB 121 LVGNKCDLTTKKVVDYTTAKEFADSLGIPIELTSAKNATNVBOSFMTMAAEIKKMGPGA 180
QY 149 TAGAEKSNVKTOSTPVKQSGGGCC 173
DB 181 TAGAEKSNVKTOSTPVKQSGGGCC 205
```